Cours de calcul différentiel
Licence de mathématiques, 3ème année

Raphaël Danchin

23 novembre 2010

Table des matières

Chapitre 1

Fonctions de plusieurs variables

1.1 Quelques notations

Dans tout ce cours, E et F sont deux $\mathbb{R}$-espaces vectoriels. Sauf mention contraire, on supposera que E et F sont *de dimension finie*, et l'on notera n la dimension de E et p la dimension de F. En pratique, on prendra très souvent $E = \mathbb{R}^n$ et $F = \mathbb{R}^p$.

Nous supposerons de plus que E (resp. F) est muni d'une norme notée $\|\cdot\|_E$ (resp. $\|\cdot\|_F$). Le lecteur n'est pas sans ignorer qu'en dimension finie toutes les normes sont équivalentes (cf le cours sur les fonctions de plusieurs variables en deuxième année de licence), et le choix de la norme n'influe donc pas sur les propriétés topologiques (continuité, limite, etc.) de E. Lorsque $E = \mathbb{R}^n$ et $x = (x_1, \cdots, x_n)$, les choix les plus courants sont

$$\|x\|_E = \sup_{i\in\{1,\cdots,n\}} |x_i|, \quad \|x\|_E = \sqrt{\sum_{i=1}^n x_i^2} \quad \text{ou} \quad \|x\|_E = \sum_{i=1}^n |x_i|.$$

Pour tout $a \in E$ et $r \geq 0$, on note $B_E(a,r)$ (resp. $\overline{B}_E(a,r)$) la **boule ouverte** (resp. la **boule fermée**) de E, c'est-à-dire l'ensemble des $x \in E$ tels que $\|x-a\|_E < r$ (resp. $\|x-a\|_E \leq r$.)

Le but premier de ce cours est d'étudier la généralisation de la notion de dérivabilité aux fonctions de *plusieurs* variables. Notre objet d'étude sera typiquement une fonction f définie sur une partie U de E (généralement supposée ouverte) et à valeurs dans F.

Si l'ensemble d'arrivée F est $\mathbb{R}^p$, on peut associer à f ses p **fonctions composantes** $f_1, \cdots, f_p$ définies par

$$f(x) = (f_1(x), \cdots, f_p(x)).$$

Dans le cas particulier où $E = F = \mathbb{R}^p$ et $f(x) = x$, la i-ème fonction composante est

$$\pi_i \;:\; \left\{ \begin{array}{ccc} \mathbb{R}^p & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x_i \end{array} \right.$$

et est appelée i-ème **projection canonique** (ou i-ème **application coordonnée**). Si $f \in \mathcal{F}(\mathbb{R}^n;\mathbb{R}^p)$, on a donc $f_i = \pi_i \circ f$.

1.2 Continuité

1.2.1 Généralités

Rappelons d'abord la définition de la continuité.

Définition 1.2.1 *Soit A une partie de E et f une application de A dans F. Soit de plus $a \in A$. On dit que f est* **continue** *en a si*

$$\boxed{\forall \epsilon > 0,\ \exists \eta > 0, \forall x \in A, \|x-a\|_E < \eta \Longrightarrow \|f(x)-f(a)\|_F < \epsilon.}$$

On dit que f est continue sur A si f est continue en tout point de A.

Exemple : On rappelle que pour $M \in \mathbb{R}^+$, une fonction $f : A \to F$ est dite **M-lipschitzienne** si elle vérifie

$$\forall (x,y) \in A^2,\ \|f(y)-f(x)\|_F \leq M\|y-x\|_E.$$

De la définition de la continuité, on déduit aisément que toute fonction lipschitzienne est continue sur son domaine de définition.

Pour montrer qu'une fonction est continue (ou qu'un ensemble est ouvert), on fait souvent appel au résultat suivant :

Proposition 1.2.2 *Soit $A \subset E$ et $f \in \mathcal{F}(A;F)$. L'application f est continue sur A si et seulement si pour tout ouvert V de F, il existe un ouvert U de E tel que $f^{-1}(V) = U \cap A$.*

Remarque : De même, l'application f est continue sur A si et seulement si pour tout fermé $\widetilde{V}$ de F, il existe un fermé $\widetilde{U}$ de E tel que $f^{-1}(\widetilde{V}) = \widetilde{U} \cap A$.

Retenons aussi le résultat suivant :

Proposition 1.2.3 *L'image d'un compact par une application continue est compacte.*

Remarque : En dimension finie, retenons que l'image d'un fermé borné par une application continue est un fermé borné.

La notion de continuité est stable par combinaison linéaire. Plus précisément, on a le résultat suivant :

Proposition 1.2.4 *Soit f et g deux fonctions de $A \subset E$ à valeurs dans F, et $a \in A$. Supposons que f et g soient continues en a. Alors pour tout couple (λ, μ) de $\mathbb{R}^2$, la fonction $\lambda f + \mu g$ est continue en a.*

Si les fonctions f et g sont continues sur A alors $\lambda f + \mu g$ aussi.

La propriété de continuité est également stable par composition :

Proposition 1.2.5 *Soit E, F, G trois espaces vectoriels de dimension finie, U une partie de E et V une partie de F. Soit $f \in \mathcal{F}(U;F)$ et $g \in \mathcal{F}(V;G)$. Soit $a \in U$. Supposons que f soit à valeurs dans V, continue en a et que g soit continue en $f(a)$. Alors $g \circ f$ est continue en a.*

Si f est à valeurs dans V et continue sur U, et si g est continue sur V alors $g \circ f$ est continue sur U.

Les quatre propositions énoncées ci-dessus sont vraies dans un cadre beaucoup plus général. Pour leur démonstration, nous renvoyons le lecteur au cours sur les fonctions de plusieurs variables, ou à [3] ou [7].

Exemples :

1. L'application $\left\{ \begin{array}{lcl} \mathrm{GL}_n(\mathbb{R}) & \longrightarrow & \mathrm{GL}_n(\mathbb{R}) \\ A & \longmapsto & A^{-1} \end{array} \right.$ est continue.
2. L'ensemble des applications continues de E dans $\mathbb{R}$ est une algèbre. De plus, si f et g sont continues en x_0 et $g(x_0) \neq 0$ alors f/g est continue en x_0.
3. Des points précédents, on déduit que toute fonction construite à partir de fonctions continues par addition, multiplication, quotient, composition, sont continues.

   Par exemple $\left\{ \begin{array}{lcl} \mathbb{R}^2 \backslash \{(0,0)\} & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & \frac{e^{\sin y}}{x^2+\pi y^2} \end{array} \right.$ est continue.

1.2.2 Applications linéaires continues

Dans toute cette partie, E et F désignent deux $\mathbb{R}$-**espaces vectoriels normés** (i.e. munis d'une norme) de dimension quelconque. On note $L(E;F)$ l'ensemble des applications linéaires de E vers F.

L'équivalence entre les deux premières assertions de la proposition suivante sera fondamentale pour la suite du cours.

Proposition 1.2.6 *Soit $u \in L(E;F)$. Les cinq propriétés suivantes sont équivalentes.*

(i) $\exists M \geq 0, \forall x \in E, \|u(x)\|_F \leq M\|x\|_E$,

(ii) *u est continue sur E,*

(iii) *u est continue en 0,*

(iv) *u est bornée sur la boule unité,*

(v) *u est bornée sur la sphère unité.*

Preuve : Il suffit de prouver $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (v) \Rightarrow (i)$.

$(i) \Rightarrow (ii)$ Par linéarité de u, on a $u(y)-u(x)=u(y-x)$. Donc,

$$\forall (x,y) \in E^2,\ \|u(y)-u(x)\|_F \leq M\|y-x\|_E.$$

Donc u est lipschitzienne donc continue.

$(ii) \Rightarrow (iii)$ Trivial.

$(iii) \Rightarrow (iv)$ De la continuité en 0 de u, on déduit l'existence d'un $\eta > 0$ tel que $\|u(x)\|_F \leq 1$ dès que $x \in \overline{B}_E(0,\eta)$. Or $x \in \overline{B}_E(0,\eta)$ si et seulement si $\eta^{-1}x \in \overline{B}_E(0,1)$. Par linéarité de u, on conclut que $\|u(x)\|_F \leq \eta^{-1}$ pour tout $x \in \overline{B}_E(0,1)$. Donc u est bornée sur la boule unité.

$(iv) \Rightarrow (v)$ Trivial.

$(v) \Rightarrow (i)$ Notons M une borne de u restreinte à la sphère unité. Si $x \neq 0$, le point $x/\|x\|_E$ appartient à la sphère unité. En utilisant la linéarité de u et le fait que u est bornée par M sur la sphère unité, on obtient donc

$$\|u(x)\|_F = \|x\|_E \left\| u\Big(\frac{x}{\|x\|_E}\Big)\right\|_F \leq M\|x\|_E.$$

■

Définition 1.2.7 *Soit $\mathcal{L}(E,F)$ l'ensemble des applications linéaires continues de E dans F. Pour $f \in \mathcal{L}(E,F)$, on note $|||f|||$ la plus petite constante M telle que*

$$\forall x \in E,\ \|f(x)\|_F \leq M\|x\|_E.$$

Proposition 1.2.8 *L'espace $(\mathcal{L}(E,F);|||\cdot|||)$ est un espace vectoriel normé.*

Preuve : Vérifions rapidement que $|||\cdot|||$ est une norme. La proposition 1.2.6 assure que pour tout $f \in \mathcal{L}(E,F)$, la quantité $|||f|||$ est finie (et positive). Il est de plus immédiat que $|||\lambda f||| = |\lambda| |||f|||$ pour tout $\lambda \in \mathbb{K}$ et que $|||f|||=0$ si et seulement si $f=0$.

Si f et g sont deux éléments de $\mathcal{L}(E,F)$, on a pour tout $x \in E$,

$$\|(f+g)(x)\|_F = \|f(x)+g(x)\|_F \leq \|f(x)\|_F + \|g(x)\|_F \leq (|||f|||+|||g|||)\|x\|_E.$$

Donc $f+g$ est linéaire continue et vérifie $|||f+g||| \leq |||f|||+|||g|||$. Donc $|||\cdot|||$ est bien une norme. ■

Nous laissons au lecteur le soin d'établir le résultat fort utile suivant :

Proposition 1.2.9 *Pour tout* $f \in \mathcal{L}(E,F)$, *on a*

$$|||f||| = \sup_{x\in E\setminus\{0\}} \frac{\|f(x)\|_F}{\|x\|_E} = \sup_{\substack{x\in E \\ \|x\|_E=1}} \|f(x)\|_F. \tag{1.1}$$

Proposition 1.2.10 *Sur* $\mathcal{L}(E) \overset{\text{déf}}{=} \mathcal{L}(E,E)$ *la norme* $|||\cdot|||$ *vérifie l'inégalité suivante :*

$$\forall u \in \mathcal{L}(E),\, \forall v \in \mathcal{L}(E),\, |||u\circ v||| \le |||u|||\,|||v|||.$$

On dit que $(\mathcal{L}(E); |||\cdot|||)$ *est une* **algèbre normée** *et que* $|||\cdot|||$ *est une* **norme d'algèbre**.

Preuve : Pour tout couple (u,v) d'éléments de $\mathcal{L}(E)$, on a

$$|||u\circ v||| = \sup_{\substack{x\in E \\ \|x\|_E=1}} \|u(v(x)\|_E \le \sup_{\substack{x\in E \\ \|x\|_E=1}} |||u|||\,\|v(x)\|_E \le |||u|||\,|||v|||,$$

d'où le résultat. ■

Théorème 1.2.11 *Si l'espace vectoriel* E *est* <u>*de dimension finie*</u> *alors toute application linéaire de* E *dans* F *est continue.*

Preuve : Fixons $(e_1,\cdots,e_n)$ une base de E. Soit $f\in\mathcal{L}(E,F)$. Comme toutes les normes sont équivalentes sur E, on peut supposer que E est muni de la norme $\|x\|_E \overset{\text{déf}}{=} \sup_{1\le i\le n}|x_i|$ où $(x_1,\cdots,x_n)$ désigne les coordonnées de x par rapport à la base $(e_1,\cdots,e_n)$.
On a alors grâce à la linéarité de f et à l'inégalité triangulaire,

$$\begin{aligned}\|f(x)\|_F &= \|\textstyle\sum_{i=1} x_i f(e_i)\|_F\,,\\ &\le \textstyle\sum_{i=1}^n |x_i|\|f(e_i)\|_F,\\ &\le (\max_{1\le i\le n}\|f(e_i)\|_F)\ \|x\|_E.\end{aligned}$$

Donc, d'après la proposition 1.2.6, l'application f est continue. ■

Exemple : Toutes les applications coordonnées π_i sont continues de $\mathbb{R}^p$ dans $\mathbb{R}$. D'après le théorème de composition, on en déduit donc que si une application $f : A\subset E \to \mathbb{R}^p$ est continue en a, il en est de même de toutes ses applications composantes.

1.2.3 Applications multilinéaires continues

Nous laissons au lecteur le soin d'établir le résultat suivant :

Proposition 1.2.12 *Soit* E_1, $\cdots$, E_p *et* F *des espaces vectoriels sur* $\mathbb{R}$, *munis de normes* $\|\cdot\|_{E_1}$, $\cdots$, $\|\cdot\|_{E_p}$ *et* $\|\cdot\|_F$. *Soit* u *une application* p*-linéaire de* $E_1\times\cdots\times E_p$ *dans* F.
Les quatre propriétés suivantes sont équivalentes.

(i) u *est continue,*
(ii) u *est continue en* $(0,\cdots,0)$,
(iii) u *est bornée sur* $B_{E_1}(0,1)\times\cdots\times B_{E_p}(0,1)$,
(iv) *Il existe* $M\in\mathbb{R}^+$ *tel que*

$$\forall (x_1,\cdots,x_p)\in E_1\times\cdots\times E_p,\ \|u(x_1,\cdots,x_p)\|_F \le M\|x_1\|_{E_1}\cdots\|x_p\|_{E_p}.$$

On note $\mathcal{L}(E_1\times\cdots\times E_p;F)$ l'ensemble des applications p-linéaires continues de $E_1\times\cdots\times E_p$ vers F.

Le résultat suivant est une généralisation du théorème 1.2.11 au cas des applications multilinéaires.

Théorème 1.2.13 *Supposons que* $E_1,\cdots,E_p$ *soient* <u>*de dimension finie*</u>. *Alors toute application multilinéaire de* $E_1\times\cdots\times E_p$ *dans* F *est continue.*

1.3 Dérivées directionnelles et dérivées partielles

Définition 1.3.1 *Soit f une application d'un ouvert U de $\mathbb{R}^n$ à valeurs dans F. Soit $i \in \{1, \cdots, n\}$ et $a \in U$. Soit $U_i^a = \{t \in \mathbb{R} \mid (a_1, \cdots, a_{i-1}, t, a_{i+1}, \cdots, a_n) \in U\}$.*

Définissons l'extension τ_i^a suivante :

$$\tau_i^a \ : \left\{ \begin{array}{lcl} U_i^a & \longrightarrow & \mathbb{R}^n \\ t & \longmapsto & (a_1, \cdots, a_{i-1}, t, a_{i+1}, \cdots, a_n). \end{array} \right.$$

L'application $f \circ \tau_i^a$ est appelée i-ème **application partielle** *de f en a.*

Remarque : La i-ème application partielle de f en a n'est autre que

$$\left\{ \begin{array}{lcl} U_i^a & \longrightarrow & F \\ t & \longmapsto & f(a_1, \cdots, a_{i-1}, t, a_{i+1}, \cdots, a_n). \end{array} \right.$$

L'ensemble U_i^a est une partie de $\mathbb{R}$ qui coïncide avec la projection sur le i-ème axe de l'intersection de l'ouvert U avec les $n-1$ hyperplans $x_j = a_j$ $(j \neq i)$. On montre facilement que U_i^a est un ouvert. De plus, les applications τ_i^a peuvent s'écrire comme somme d'une application constante et d'une application linéaire continue, et sont donc continues de U_i^a dans $\mathbb{R}^n$.

Si f est à valeurs réelles, le graphe de la i-ème application partielle de f en a s'obtient en faisant l'intersection du graphe de f avec tous les hyperplans $x_j = a_j$ $(j \neq i)$.

Proposition 1.3.2 *Si f est continue, toutes les applications partielles qui lui sont associées le sont également.*

Exercice : Prouver la proposition ci-dessus.

Attention : La réciproque est fausse en général. On pourra étudier la continuité en 0 de l'application

$$(x, y) \mapsto \left\{ \begin{array}{ll} 0 \quad \text{si} & (x, y) = (0, 0), \\ \frac{xy}{x^2+y^2} \quad \text{si} & (x, y) \neq (0, 0) \end{array} \right.$$

pour s'en convaincre.

Définition 1.3.3 *Soit f une application d'un ouvert U de E à valeurs dans F, v un élément de E et $a \in U$. On dit que f admet une* **dérivée en** a **suivant le vecteur** v *si la limite suivante existe :*

$$\frac{\partial f}{\partial v}(a) \stackrel{\text{déf}}{=} \lim_{t \to 0} \frac{f(a + tv) - f(a)}{t}.$$

On dit que f est **Gâteaux-différentiable** *en a si les dérivées directionnelles $\frac{\partial f}{\partial v}(a)$ sont définies pour tout $v \in E$ et si de plus l'application $v \mapsto \frac{\partial f}{\partial v}(a)$ est linéaire.*

Définition 1.3.4 *Soit f une fonction définie sur un ouvert U de $\mathbb{R}^n$ et à valeurs dans $\mathbb{R}^p$. Si f admet en a une dérivée suivant le i-ème vecteur ϵ_i de la base canonique de $\mathbb{R}^n$, on dit que f* **admet une dérivée partielle dans la direction** x_i *et l'on note $\frac{\partial f}{\partial x_i}(a)$ la dérivée de f en a suivant le vecteur ϵ_i.*

Remarques :

(1) Lorsque f ne dépend que de 2 ou 3 variables, on utilise souvent la notation (x, y) (resp. (x, y, z)) au lieu de (x_1, x_2) (resp. (x_1, x_2, x_3)) et les dérivées partielles sont notées $\frac{\partial f}{\partial x}(a)$, $\frac{\partial f}{\partial y}(a)$, etc.

(2) Le vecteur $\frac{\partial f}{\partial x_i}(a)$ n'est autre que la dérivée en a_i de la i-ème application partielle de f en a.

(3) Lorsque f est à valeurs réelles, on peut facilement donner une interprétation géométrique de $\frac{\partial f}{\partial x_i}(a)$: c'est la pente de la tangente en $(a, f(a))$ à la courbe obtenue en faisant l'intersection des hyperplans $x_j = a_j$ $(j \neq i)$ avec le graphe de f.

Si f admet une dérivée partielle selon la variable x_i, il est clair qu'il en est de même de toutes ses fonctions composantes, ce qui justifie la définition suivante.

Définition 1.3.5 *Soit U un ouvert de $\mathbb{R}^n$, $a \in U$ et f une application de U dans $\mathbb{R}^p$. Supposons que toutes les dérivées partielles de f en a soient définies et notons $f_j \stackrel{\text{déf}}{=} \pi_j \circ f$ la j-ème application composante de f. Alors tous les f_j admettent des dérivées partielles en a suivant toutes les directions données par la base canonique, et la matrice suivante de $\mathcal{M}_{p,n}(\mathbb{R})$:*

$$Df(a) \stackrel{\text{déf}}{=} \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \cdots & \frac{\partial f_2}{\partial x_n}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \frac{\partial f_p}{\partial x_2}(a) & \cdots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix}$$

est appelée **matrice jacobienne** *de f en a.*

Si $p = n$, le déterminant de la matrice jacobienne est appelé **jacobien**. *On le note $\mathrm{Jac}_f(a)$ ou $\frac{D(f_1,\cdots,f_n)}{D(x_1,\cdots,x_n)}(a)$. On a donc*

$$\mathrm{Jac}_f(a) = \frac{D(f_1, \cdots, f_n)}{D(x_1, \cdots, x_n)}(a) = \det Df(a).$$

Exemples :

(1) Pour une fonction f de $\mathbb{R}$ dans $\mathbb{R}$, la matrice jacobienne se réduit à une matrice carrée de taille 1 que l'on peut identifier à la dérivée $f'(a)$.

(2) Pour une fonction f de $\mathbb{R}^n$ dans $\mathbb{R}$, la matrice jacobienne est un vecteur ligne. La transposée de ce vecteur ligne est un vecteur colonne appelé **gradient** de f. On le note ∇f.

(3) Donnons un exemple concret de calcul de matrice jacobienne.
Soit $f : (x,y) \mapsto (\sin(x^2+y^2), xy)$. La fonction f admet des dérivées partielles suivant x et y pour tout $(x,y) \in \mathbb{R}^2$ et l'on a :

$$Df(x,y) = \begin{pmatrix} 2x\cos(x^2+y^2) & 2y\cos(x^2+y^2) \\ y & x \end{pmatrix} \quad \text{et} \quad \frac{D(f_1, f_2)}{D(x,y)} = 2(x^2-y^2)\cos(x^2+y^2).$$

Chapitre 2

Différentiabilité

Ici encore, E désigne un $\mathbb{R}$-espace vectoriel de dimension n et F, un $\mathbb{R}$-espace vectoriel de dimension p.

2.1 Définition de la différentiabilité

Rappel : Soit f une fonction d'un intervalle ouvert I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$, et a un point de I. On dit que f est dérivable en a s'il existe un nombre $f'(a)$ tel que

$$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} = f'(a). \tag{2.1}$$

Sauf si $n=1$, donner un sens à la limite ci-dessus pour une fonction allant d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$ paraît délicat.

L'idée intuitive à conserver de la notion de dérivée est celle de *meilleure approximation.*

La relation (2.1) stipule que pour une fonction f de $\mathbb{R}$ dans $\mathbb{R}$ dérivable en a, la fonction $x \mapsto f(a)+(x-a)f'(a)$ est la meilleure approximation affine de f :

$$\lim_{x\longrightarrow a} \frac{f(x)-\big(f(a)+(x-a)f'(a)\big)}{x-a} = 0.$$

C'est le bon point de vue à considérer pour pouvoir généraliser la notion de dérivabilité (que l'on appellera plutôt différentiabilité) au cas d'une fonction de n variables ou plus généralement au cas d'une fonction d'un ouvert U de E et à valeurs dans F [1].

Proposition 2.1.1 *Soit $f \in \mathcal{F}(U,F)$ avec U ouvert de E, et a un point de U. On dit que f est* **différentiable** *en a s'il existe une application linéaire $L \in \mathcal{L}(E;F)$ telle que*

$$\lim_{x\longrightarrow a} \frac{\|f(x)-\big(f(a)+L(x-a)\big)\|_F}{\|x-a\|_E} = 0. \tag{2.2}$$

L'application L est alors unique et est appelée **différentielle** *de f en a. On la note $df(a)$.*

Preuve : Supposons que f admette deux différentielles L et L' en a. Alors pour tout $\epsilon > 0$, il existe $\eta > 0$ tel que pour $h \in B(0,\eta)$, on ait

$$\|f(a+h)-f(a)-L(h)\|_F \le \epsilon\|h\|_E \quad \text{et} \quad \|f(a+h)-f(a)-L'(h)\|_F \le \epsilon\|h\|_E.$$

[1] Dans tout ce cours de licence, on se limitera à des e.v. de *dimension finie.* De ce fait, il suffit de demander à la différentielle définie ci-dessous d'être linéaire, le caractère continu étant automatiquement assuré pas la définition finie. Mais la plupart des notions introduites ici peuvent se généraliser à des espaces vectoriels normés généraux. Le lecteur intéressé pourra consulter les ouvrages cités dans la bibliographie.

On en déduit que

$$\forall h \in B(0,\eta), \|(L-L')(h)\|_F \leq 2\epsilon \|h\|_E.$$

Comme $L - L'$ est linéaire, la relation ci-dessus est en fait vraie pour tout $h \in E$.
Il ne reste plus qu'à faire tendre ϵ vers 0 pour conclure que $L = L'$. ∎

Remarques :

(1) La relation (2.2) se récrit souvent

$$f(a+h) = f(a) + df(a)(h) + o_F(h)$$

où o_F désigne une fonction définie sur un voisinage de 0, à valeurs dans F et telle que

$$\lim_{h \to 0} \frac{o_F(h)}{\|h\|_E} = 0.$$

(2) L'image du vecteur $h \in E$ par $df(a)$ est notée $df(a)(h)$ ou parfois $df(a) \cdot h$.

(3) Si f est différentiable en tout point de U, on peut considérer l'application

$$df \ : \left\{ \begin{array}{ccc} U & \longmapsto & \mathcal{L}(E;F) \\ x & \longmapsto & df(x). \end{array} \right.$$

C'est donc une fonction de $\mathcal{F}(U;\mathcal{L}(E;F))$ dont on peut étudier les propriétés de continuité. Si df est continue sur U, on dit que f est **continûment différentiable** sur U ou encore que f est **de classe** C^1 sur U.

Exemples :

(1) Différentielle d'une application linéaire. Soit $f \in \mathcal{L}(E;F)$. Alors pour tout $a \in E$, pour tout $h \in E$, on a

$$f(a+h) - f(a) = f(h).$$

Donc $\boxed{df(a)(h) = f(h).}$

En particulier, comme chaque application π_i est une forme linéaire sur $\mathbb{R}^p$, on a $d\pi_i = \pi_i$.

(2) Différentielle d'une application affine. Si $f \in \mathcal{F}(E;F)$ est de la forme $f(x) = y_0 + L(x)$ avec $y_0 \in F$ et $L \in \mathcal{L}(E;F)$, alors on a $df(a) = L$ pour tout $a \in E$.

En appliquant ce résultat à

$$\tau_j^a \ : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R}^n \\ t & \longmapsto & (0,\cdots,0,\overbrace{t}^{j},0,\cdots,0) \end{array} \right.$$

on obtient donc $d\tau_j^a = \tau_j^0$.

(3) Différentielle d'une application bilinéaire. Soit E_1, E_2 et F trois espaces vectoriels de dimension finie et B une application bilinéaire de $E_1 \times E_2$ dans F. Alors on a

$$\forall (x,y) \in E_1 \times E_2,\ \forall (h,k) \in E_1 \times E_2,\ B(x+h,y+k) - B(x,y) - B(h,y) - B(x,k) = B(h,k).$$

Comme B est continue (nous sommes en dimension finie), il existe $M \geq 0$ tel que $\|B(h,k)\|_G \leq M \|h\|_{E_1} \|k\|_{E_2}$. Autrement dit, $\|B(h,k)\|_G = O(\|(h,k)\|^2_{E_1 \times E_2})$.
Enfin, l'application $(h,k) \mapsto B(h,y) + B(x,k)$ est linéaire de $E_1 \times E_2$ dans F. Donc B est différentiable sur $E_1 \times E_2$ et

$$\boxed{\forall (h,k) \in E_1 \times E_2,\ dB(x,y)(h,k) = B(x,k) + B(h,y).}$$

Exercice : Soit $E_1, \cdots, E_k$ et F des espaces vectoriels de dimension finie et $G : E_1 \times \cdots \times E_k \to F$ une application k-linéaire. Montrer que G est différentiable en tout point et que

$$\forall (x_1, \cdots, x_k) \in E_1 \times \cdots \times E_k,\ \forall (h_1, \cdots, h_k) \in E_1 \times \cdots \times E_k,$$
$$dG(x_1, \cdots, x_k)(h_1, \cdots, h_k) = G(h_1, x_2, \cdots, x_k) + G(x_1, h_2, \cdots, x_k) + \cdots + G(x_1, \cdots, x_{k-1}, h_k).$$

Proposition 2.1.2 *Soit $f \in \mathcal{F}(U, F)$ avec U ouvert de E, et a un point de U. On a*

(i) f différentiable en a $\Longrightarrow$ f continue en a.

(ii) f différentiable en a $\Longrightarrow$ f Gâteaux-différentiable en a. De plus, on a

$$\forall v \in E,\ \frac{\partial f}{\partial v}(a) = df(a)(v).$$

Preuve : Si f est différentiable en a alors il existe une fonction ε définie sur un voisinage V de 0, à valeurs dans F et tendant vers 0 en 0 telle que

$$\forall h \in V,\ f(a+h) - f(a) - df(a)(h) = \|h\|_E\, \varepsilon(h).$$

On en déduit que $h \mapsto f(a+h) - f(a) - df(a)(h)$ est continue en 0. Comme $df(a)$ est elle-même une application continue et nulle en 0 (car linéaire), on conclut que $f(a+h) - f(a)$ tend vers 0 quand h tend vers 0, ce qui entraîne la continuité de f en a.

Pour montrer le deuxième point, il suffit de prendre h de la forme tv dans la définition de la différentiabilité et l'on trouve $\frac{\partial f}{\partial v}(a) = df(a)(v)$ ce qui assure bien la linéarité de l'application $v \mapsto \frac{\partial f}{\partial v}(a)$. ■

En appliquant le point (ii) au cas où f est une fonction de $\mathbb{R}^n$ dans $\mathbb{R}^p$, on obtient le résultat suivant :

Proposition 2.1.3 *Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ une application différentiable en a. Alors toutes les dérivées partielles $\frac{\partial f_i}{\partial x_j}(a)$ de f sont définies et l'on a pour tout $h \in \mathbb{R}^n$ (en identifiant h avec la matrice colonne de mêmes composantes) :*

$$df(a)(h) = Df(a) \cdot h.$$

Autrement dit, la matrice de $df(a)$ par rapport aux bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$ est la matrice jacobienne $Df(a)$.

Remarque : Pour une fonction de $\mathbb{R}$ dans $\mathbb{R}^p$, la proposition ci-dessus permet de faire le lien entre les notions de dérivabilité et de différentiabilité. Plus précisément, une fonction $f : \mathbb{R} \to \mathbb{R}^p$ est *dérivable* en a si et seulement si elle est *différentiable* en a, et l'on a alors

$$\forall h \in \mathbb{R},\ df(a)(h) = hf'(a).$$

Attention : Les réciproques des deux implications de la proposition 2.1.2 sont fausses en général. En particulier, ce n'est pas parce que toutes les dérivées partielles de f en a sont définies que f est automatiquement différentiable en a. Pour s'en convaincre, le lecteur pourra étudier la différentiabilité de la fonction

$$f\ :\ \begin{cases} \mathbb{R}^2 & \longrightarrow\ \mathbb{R} \\ (x,y) & \longmapsto\ \begin{cases} \frac{xy}{x^2+y^2} & \text{si}\ \ (x,y) \neq (0,0), \\ 0 & \text{si}\ \ (x,y) \neq (0,0). \end{cases} \end{cases}$$

Au chapitre 3, nous donnerons des conditions suffisantes sur les dérivées partielles pour assurer la différentiabilité.

2.2 Propriétés classiques

Proposition 2.2.1 *Soit U un ouvert de E et f, g deux fonctions de $\mathcal{F}(U;F)$ différentiables en un point a de U. Alors pour tout $(\lambda,\mu)\in\mathbb{R}^2$, la fonction $\lambda f+\mu g$ est différentiable en a et l'on a*

$$d(\lambda f+\mu g)(a)=\lambda df(a)+\mu dg(a).$$

Preuve : En revenant à la définition de la différentiabilité de f et de g en a, on constate que

$$\lim_{x\longrightarrow a}\frac{\|\big(\lambda f(x)+\mu g(x)\big)-\big(\lambda f(a)+\mu g(a)\big)-\big(\lambda df(a)(x-a)+\mu dg(a)(x-a)\big)\|_F}{\|x-a\|_E}=0.$$

Comme l'application $h\mapsto \lambda df(a)(h)+\mu dg(a)(h)$ est linéaire, on obtient bien le résultat souhaité. ∎

Théorème 2.2.2 (de composition) *Soit E, F, G trois espaces vectoriels de dimension finie. Soit U un ouvert de E et V un ouvert de F. Soit enfin $a\in U$, $f\in\mathcal{F}(U;V)$ et $g\in\mathcal{F}(V;G)$. Supposons que f soit différentiable en a et que g soit différentiable en $f(a)$. Alors $g\circ f$ est différentiable en a et l'on a*

$$\boxed{d(g\circ f)(a)=dg(f(a))\circ df(a).}$$

Preuve : Comme f est différentiable en a, on a

$$f(a+h)-f(a)=df(a)(h)+o_F(h).$$

De même, comme g est différentiable en $f(a)$, on a

$$g(f(a)+k)-g(f(a))=dg(f(a))(k)+o_G(k).$$

Si l'on choisit justement $k=f(a+h)-f(a)$, on obtient donc compte tenu du fait que $dg(f(a))$ et $df(a)$ sont des applications linéaires continues :

$$\begin{aligned} g(f(a+h))-g(f(a)) &= dg(f(a))(f(a+h)-f(a))+o_G(f(a+h)-f(a)),\\ &= dg(f(a))(df(a)(h))+o_G(h). \end{aligned}$$

∎

Exemples :

(1) Soit $f:U\subset\mathbb{R}^n\mapsto\mathbb{R}^p$ différentiable en $a\in U$. Soit $h>0$ et γ une application de $]-h,h[$ à valeurs dans U dérivable en 0 et telle que $\gamma(0)=a$ [2]. Notons $v=\gamma'(0)$.
Alors le théorème de composition assure que $f\circ\gamma$ est dérivable en 0 et que l'on a

$$\boxed{(f\circ\gamma)'(0)=df(a)(\gamma'(0))=\frac{\partial f}{\partial v}(a).}$$

Si l'on choisit $\gamma(t)=a+tv$, on retrouve la définition de la dérivée suivant le vecteur v.

(2) Soit E, F, G et H quatre espaces vectoriels de dimension finie, U un ouvert de E et $a\in U$. Soit B une application bilinéaire de $F\times G$ dans H, $f\in\mathcal{F}(U;F)$ et $g\in\mathcal{F}(U;G)$. Si f et g sont différentiables en a alors l'application $\Psi : x\mapsto B(f(x),g(x))$ aussi et l'on a

$$\begin{aligned} \forall h\in E,\ d\Psi(a)(h) &= dB(f(a),g(a))\big(df(a)(h),dg(a)(h)\big),\\ &= B\big(df(a)(h),g(a)\big)+B\big(f(a),dg(a)(h)\big). \end{aligned}$$

[2] Autrement dit, le graphe de γ est une courbe de $\mathbb{R}^n$ passant par le point a.

Dans le cas où $F = G = H$, si l'on note '$*$' l'application bilinéaire B, on a donc la formule suivante

$$\boxed{\forall h \in E,\ d(f * g)(a)(h) = (df(a)(h)) * g(a) + f(a) * (dg(a)(h))} \tag{2.3}$$

qui peut être vue comme une généralisation de la formule de Leibniz $(fg)' = f'g + fg'$ pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$.

Les exemples d'applications de la formule (2.3) sont nombreux : multiplication de matrices, produit scalaire, etc. On peut aisément la généraliser au cas multilinéaire.

Corollaire 2.2.3 *Soit U un ouvert de E. La fonction $f : U \to \mathbb{R}^p$ est différentiable en $a \in U$ si et seulement si toutes ses applications composantes f_i le sont. On a alors*

$$\forall h \in E,\ df(a)(h) = (df_1(a)(h), \cdots, df_p(a)(h)).$$

Preuve : On utilise $f_i = \pi_i \circ f$ et le théorème de composition. ∎

Corollaire 2.2.4 *Soit U un ouvert de $\mathbb{R}^n$ et V un ouvert de $\mathbb{R}^p$. Soit $f : U \to V$ différentiable en $a \in U$ et $g : V \to \mathbb{R}^q$ différentiable en $f(a)$. Alors $g \circ f$ est différentiable en a et l'on a, au sens de la multiplication des matrices de $\mathcal{M}_{q,p}(\mathbb{R})$ par les matrices de $\mathcal{M}_{p,n}(\mathbb{R})$:*

$$\boxed{D(g \circ f)(a) = Dg(f(a))\, Df(a).}$$

Autrement dit,

$$\boxed{\forall j \in \{1, \cdots, n\},\ \forall i \in \{1, \cdots, q\},\ \frac{\partial (g \circ f)_i}{\partial x_j}(a) = \sum_{k=1}^{p} \frac{\partial g_i}{\partial y_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a).}$$

Preuve : Il suffit d'utiliser le fait que la matrice jacobienne d'une application différentiable n'est autre que la matrice de sa différentielle par rapport à la base canonique, et d'appliquer le théorème de composition. ∎

2.3 La notation différentielle et les changements de variables

Dans toute cette section, U désigne un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}$. Tous les résultats qui suivent se généralisent sans peine au cas d'une fonction à valeurs *vectorielles* : il suffit de les appliquer à chaque composante d'une telle application.

Si l'on suppose f différentiable en a, nous avons vu que

$$\forall h \in \mathbb{R}^n,\ df(a)(h) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) h_i = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\, \pi_i(h). \tag{2.4}$$

En calcul différentiel, il est d'usage de noter dx_i l'application π_i et l'on obtient donc le résultat suivant :

Proposition 2.3.1 *Soit $f : U \to \mathbb{R}$ différentiable en a. Alors on a la formule*

$$\boxed{df(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\, dx_i.} \tag{2.5}$$

La “popularité” de cette formule (que vous avez certainement rencontrée dans les cours de thermodynamique) tient à son “invariance par changement de variable”.

En effet, considérons un ouvert V de $\mathbb{R}^p$ et $\varphi : V \to U$ différentiable sur V. Supposons $f : U \to \mathbb{R}$ différentiable sur U. Pour $t \in V$, notons $t = (t_1, \cdots, t_p)$ et $\frac{\partial \varphi}{\partial t_i}$ les dérivées partielles de φ. De même pour $x \in U$, notons $x = (x_1, \cdots, x_n)$ et $\frac{\partial f}{\partial x_j}$ les dérivées partielles de f.

D'après le théorème de composition, la fonction $F \overset{\text{déf}}{=} f \circ \varphi$ est différentiable sur V et l'on a

$$\forall i \in \{1, \cdots, p\},\ \forall t \in V,\ \frac{\partial F}{\partial t_i}(t) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\varphi(t)) \frac{\partial \varphi_j}{\partial t_i}(t).$$

En notant dt_i la i-ème application coordonnée sur $\mathbb{R}^p$, on a donc

$$dF(t) = \sum_{i=1}^{p} \frac{\partial F}{\partial t_i}(t)\, dt_i = \sum_{i=1}^{p} \biggl(\sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\varphi(t)) \frac{\partial \varphi_j}{\partial t_i}(t) \biggr) dt_i = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\varphi(t)) \biggl(\sum_{i=1}^{p} \frac{\partial \varphi_j}{\partial t_i}(t)\, dt_i \biggr).$$

Mais la somme en i se trouvant dans le dernier terme n'est autre que la différentielle de φ_j. Donc on a

$$\boxed{dF(t) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\varphi(t))\, d\varphi_j(t).} \tag{2.6}$$

On commet souvent l'abus de notation consistant à donner le même nom à la fonction f et à la fonction $f \circ \varphi$ en écrivant $f(x)$ pour parler de f et $f(t)$ pour parler de F. Pire, on écrit $x_j = x_j(t)$ pour parler de la fonction φ_j !

Cet abus est justifié dans la mesure où la formule (2.6) garantit que les dx_i dans (2.5) se comportent bien comme des différentielles et permet de gagner du temps dans les calculs. Néanmoins, cette convention, couramment employée en physique et en mathématiques appliquées peut être fort dangereuse et doit être utilisée avec précaution... En cas de doute, il est conseillé de se reporter aux formules du corollaire 2.2.4.

Exemple : Coordonnées polaires. Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction différentiable et $\varphi : (\rho, \theta) \mapsto (\rho\cos\theta, \rho\sin\theta)$ définie sur $\mathbb{R} \setminus \{0\} \times \mathbb{R}$. On cherche à calculer la différentielle de $F \overset{\text{déf}}{=} f \circ \varphi$.

Méthode matricielle :

On calcule la matrice jacobienne de φ (qui est une matrice 2×2) et la matrice jacobienne de f (c'est une matrice ligne à deux composantes) et on applique le théorème de composition pour déterminer la matrice jacobienne de F.

Méthode différentielle :

On part de l'expression de la différentielle de f :

$$df = \frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial y}\, dy \tag{2.7}$$

et l'on veut calculer $\frac{\partial F}{\partial \rho}$ et $\frac{\partial F}{\partial \theta}$ tels que la relation $dF = \frac{\partial F}{\partial \rho}\, d\rho + \frac{\partial F}{\partial \theta}\, d\theta$ soit vérifiée. Notons (x, y) les composantes de φ. on a

$$dx = \frac{\partial x}{\partial \rho}\, d\rho + \frac{\partial x}{\partial \theta}\, d\theta = \cos\theta\, d\rho - \rho\sin\theta\, d\theta \quad \text{et} \quad dy = \frac{\partial y}{\partial \rho}\, d\rho + \frac{\partial y}{\partial \theta}\, d\theta = \sin\theta\, d\rho + \rho\cos\theta\, d\theta.$$

En reportant ces deux égalités dans (2.7), on obtient finalement

$$\begin{aligned} dF(\rho,\theta) &= \tfrac{\partial f}{\partial x}(x,y)\,(\cos\theta\, d\rho - \rho\sin\theta\, d\theta) + \tfrac{\partial f}{\partial y}(x,y)\,(\sin\theta\, d\rho + \rho\cos\theta\, d\theta)\,, \\ &= \underbrace{\Bigl(\cos\theta\, \tfrac{\partial f}{\partial x}(x,y) + \sin\theta\, \tfrac{\partial f}{\partial y}(x,y)\Bigr)}_{\frac{\partial F}{\partial \rho}(\rho,\theta)} d\rho + \underbrace{\rho\Bigl(\cos\theta\, \tfrac{\partial f}{\partial y}(x,y) - \sin\theta\, \tfrac{\partial f}{\partial x}(x,y)\Bigr)}_{\frac{\partial F}{\partial \theta}(\rho,\theta)} d\theta. \end{aligned}$$

Chapitre 3

L'inégalité des accroissements finis

3.1 Le cas d'une fonction numérique d'une variable réelle

Pour une fonction de $\mathbb{R}$ dans $\mathbb{R}$, nous disposons du résultat classique suivant apppelé **égalité des accroissements finis** :

Théorème 3.1.1 *Soit f une fonction continue de $[a,b]$ dans $\mathbb{R}$, et dérivable sur $]a,b[$. Alors il existe $c\in]a,b[$ tel que*

$$f(b)-f(a)=(b-a)f'(c).$$

Preuve : On pose $\varphi(t)=f(t)-f(a)-k(t-a)$ où k est choisi de telle sorte que $\varphi(b)=0$. Comme on a bien sûr $\varphi(a)=0$ (quel que soit le choix de k), on peut appliquer le théorème de Rolle. On en déduit l'existence d'un $c\in]a,b[$ tel que $\varphi'(c)=0$. Un calcul immédiat montre que $\varphi'(c)=0$ si et seulement si $k=f'(c)$. Comme par ailleurs

$$k=\frac{f(b)-f(a)}{b-a},$$

on obtient bien le résultat voulu. ■

Théorème 3.1.2 *Soit f une fonction continue d'un intervalle ouvert I de $\mathbb{R}$ et à valeurs dans $\mathbb{R}$. Supposons f dérivable sur I à dérivée bornée et notons $M=\sup_{t\in I}|f'(t)|$. Alors on a*

$$\forall(x,y)\in I^2,\ |f(y)-f(x)|\leq M|x-y|. \qquad (3.1)$$

Preuve : Soit $(x,y)\in I^2$. L'égalité des accroissements finis assure l'existence d'un $z\in I$ tel que

$$f(y)-f(x)=(y-x)f'(z).$$

Pour conclure, il suffit de prendre la valeur absolue des deux membres de l'égalité et d'utiliser le fait que $|f'(z)|\leq M$. ■

3.2 Le théorème des accroissements finis pour les fonctions de plusieurs variables

Commençons par rappeler la définition d'un segment d'un espace vectoriel réel.

Définition 3.2.1 *Soit E un e.v. sur $\mathbb{R}$ et $(a,b)\in E^2$. On appelle* **segment fermé** *d'extrémités a et b l'ensemble*

$$[a,b]\stackrel{\text{déf}}{=}\{z\in E\ \mid\ \exists t\in[0,1],\ z=ta+(1-t)b\}.$$

On appelle **segment ouvert** *d'extrémités* a *et* b *l'ensemble*

$$]a,b[\stackrel{\text{déf}}{=} \{z \in E \mid \exists t \in]0,1[,\ z = ta + (1-t)b\}.$$

L'égalité des accroissements finis se généralise facilement aux fonctions de plusieurs variables *à valeurs réelles* :

Théorème 3.2.2 *Soit f une fonction continue sur un ouvert U de E et à valeurs réelles. Soit $(a,b) \in U^2$ tel que $[a,b] \subset U$. Supposons f différentiable en tout point de $]a,b[$. Alors il existe $c \in]a,b[$ tel que*

$$f(b) - f(a) = df(c)(b-a).$$

Preuve : Pour $t \in [0,1]$, on pose $\varphi(t) = f(a + t(b-a))$. Grâce au théorème de composition, φ est continue sur $[0,1]$, dérivable sur $]0,1[$ et $\varphi'(t) = df(a + t(b-a))(b-a)$. L'égalité des accroissements finis "classique" s'applique donc à φ et donne le résultat voulu. ■

Considérons l'application suivante :

$$\varphi \ : \ \begin{cases} \mathbb{R} & \longrightarrow \quad \mathbb{R}^2 \\ t & \longmapsto \quad (\cos t, \sin t). \end{cases}$$

Il est clair que $\varphi(0) = \varphi(2\pi)$. Pourtant φ' ne s'annule jamais sur $[0, 2\pi]$.

Cet exemple simple montre que pour les fonctions *à valeurs vectorielles*, il n'y a aucun espoir d'avoir une égalité des accroissements finis similaire à celle du cas réel.

En revanche, l'inégalité des accroissements finis demeure vraie. Avant de l'énoncer dans le cas général, rappelons que la norme d'une application linéaire continue de E vers F est définie par

$$|||u||| \stackrel{\text{déf}}{=} \sup_{\|x\|_E = 1} \|u(x)\|_F = \sup_{x \neq 0} \frac{\|u(x)\|_F}{\|x\|_E}.$$

L'inégalité des accroissements finis pour les fonctions de E dans F peut alors se généraliser ainsi :

Théorème 3.2.3 *Soit f une fonction continue sur un ouvert U de E et à valeurs dans F. Supposons f différentiable sur le segment ouvert $]a,b[$ de U et*

$$M \stackrel{\text{déf}}{=} \sup_{t \in]0,1[} |||df(a + t(b-a))||| < +\infty.$$

Alors on a l'inégalité suivante :

$$\|f(b) - f(a)\|_F \leq M \|b - a\|_E.$$

Preuve : Soit f vérifiant les hypothèses de l'énoncé. Supposons de plus $a \neq b$ (sinon le résultat est trivial) et f différentiable en a avec $|||df(a)||| \leq M$. Fixons $\epsilon > 0$. Soit

$$A_\epsilon \stackrel{\text{déf}}{=} \big\{t \in [0,1] \mid \|f(a + t(b-a)) - f(a)\|_F \leq (M + \epsilon) t \|b - a\|_E\big\}.$$

Pour $t \in [0,1]$, notons $g_\epsilon(t) = (M + \epsilon)t\|b - a\|_E - \|f(a + t(b-a)) - f(a)\|_F$.
Remarquons que $A_\epsilon = g_\epsilon^{-1}(\mathbb{R}^+)$. Comme f est continue sur U, l'application g_ϵ est continue sur $[0,1]$. L'ensemble A_ϵ est donc un fermé de $[0,1]$. Il est non vide car contient 0, et évidemment borné. Donc A_ϵ admet un plus grand élément t_0.
On veut montrer que $t_0 = 1$. Supposons par l'absurde que $t_0 < 1$. Alors on a

$$\forall t \in]t_0, 1],\ \|f(a + t(b-a)) - f(a)\|_F > (M + \epsilon)t\|b - a\|_E.$$

En soustrayant l'inégalité

$$\|f(a+t_0(b-a))-f(a)\|_F \leq (M+\epsilon)t_0\|b-a\|_E,$$

et en appliquant la deuxième inégalité triangulaire, on conclut que

$$(3.2) \qquad \forall t \in]t_0,1],\ \|f(a+t(b-a))-f(a+t_0(b-a))\|_F > (M+\epsilon)(t-t_0)\|b-a\|_E.$$

Or

$$\lim_{t\to t_0} \frac{f(a+t(b-a))-f(a+t_0(b-a))}{t-t_0} = df(a+t_0(b-a))(b-a).$$

En vertu de (3.2), on a donc $|||df(a+t_0(b-a))||| \geq M+\epsilon$, ce qui est contraire aux hypothèses. On conclut que $1 \in A_\epsilon$ et il ne reste plus qu'à faire tendre ϵ vers 0 pour obtenir le résultat voulu. ■

Remarque : Le cas où f n'est pas différentiable en a se traite en appliquant le résultat précédent sur le segment $[a+\eta(b-a),b]$ (avec le même M) puis en faisant tendre η vers 0.

3.3 Quelques applications de l'inégalité des accroissements finis

En pratique, il est souvent plus facile de calculer des dérivées partielles que d'établir la différentiabilité. Nous avons déjà souligné le fait que l'existence de toutes les dérivées partielles en un point n'entraînait pas forcément la différentiabilité. Nous disposons heureusement du résultat suivant qui permet de conclure à la différentiabilité dans la plupart des cas :

Théorème 3.3.1 *Soit $f \in \mathcal{F}(U,\mathbb{R}^p)$ avec U ouvert de $\mathbb{R}^n$ et a un point de U. Supposons que toutes les dérivées partielles de f existent au voisinage de a et soient continues en a. Alors f est différentiable en a et l'on a*

$$\forall h \in E,\ df(a)(h) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a)h_i.$$

Preuve : Soit $h=(h_1,\cdots,h_n)\in\mathbb{R}^n$. on a

$$\begin{aligned} f(a+h)-f(a)-\sum_{i=1}^n \frac{\partial f}{\partial x_i}(a)h_i &= \Big[f(a_1+h_1,a_2,\cdots,a_n)-f(a_1,a_2,\cdots,a_n)-h_1\frac{\partial f}{\partial x_1}(a)\Big] \\ &+\Big[f(a_1+h_1,a_2+h_2,a_3,\cdots,a_n)-f(a_1+h_1,a_2,\cdots,a_n)-h_2\frac{\partial f}{\partial x_2}(a)\Big]+\cdots \\ &\cdots+\Big[f(a_1+h_1,\cdots,a_n+h_n)-f(a_1+h_1,\cdots,a_{n-1}+h_{n-1},a_n)-h_n\frac{\partial f}{\partial x_n}(a)\Big]. \end{aligned}$$

Fixons un $\epsilon>0$. Remarquons que $t\mapsto f(t,a_2,\cdots,a_n)$ n'est autre que la première application partielle associée à f en a qui, par hypothèse, est dérivable en a_1. Par conséquent, il existe $\eta_1>0$ tel que

$$\forall h_1 \in]-\eta_1,\eta_1[,\ \Big\|f(a_1+h_1,a_2,\cdots,a_n)-f(a_1,a_2,\cdots,a_n)-h_1\frac{\partial f}{dx_1}(a)\Big\|_{\mathbb{R}^p} \leq \epsilon\,|h_1|.$$

Remarquons ensuite que le deuxième terme entre crochets peut se récrire $\varphi_{h_1}(h_2)-\varphi_{h_1}(0)$ avec

$$\varphi_{h_1}(t) \overset{\text{déf}}{=} f(a_1+h_1,a_2+t,a_3,\cdots,a_n)-t\frac{\partial f}{\partial x_2}(a).$$

La fonction φ_{h_1} est dérivable près de 0, et l'on a pour t assez petit

$$\varphi'_{h_1}(t) = \frac{\partial f}{\partial x_2}(a_1 + h_1, a_2 + t, a_3, \cdots, a_n) - \frac{\partial f}{\partial x_2}(a).$$

Grâce à la continuité des dérivées partielles en 0, il existe donc $\eta_2 > 0$ tel que pour tout $(h_1, h_2) \in]-\eta_2, \eta_2[^2$ on ait $\|\varphi'(t)\|_F \leq \epsilon$.
En vertu du théorème des accroissements finis, on a donc pour tout $(h_1, h_2) \in]-\eta_2, \eta_2[^2$,

$$\left\| f(a_1+h_1, a_2+h_2, a_3, \cdots, a_n) - f(a_1+h_1, a_2, \cdots, a_n) - h_2 \frac{\partial f}{\partial x_2}(a) \right\|_{\mathbb{R}^p} \leq \epsilon\, |h_2|.$$

Les termes suivants se traitent de même. En posant $\eta = \min(\eta_1, \cdots, \eta_n)$, on obtient finalement

$$\forall h \in \mathbb{R}^n,\ \max_{i \in \{1,\cdots,n\}} |h_i| < \eta \Longrightarrow \left\| f(a+h) - f(a) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) h_i \right\|_{\mathbb{R}^p} \leq \epsilon \sum_{i=1}^{n} |h_i|,$$

d'où le résultat. ∎

Remarque : Si les hypothèses du théorème 3.3.1 sont vérifiées, on dit que f est **continûment différentiable** en a.

Corollaire 3.3.2 *Soit U un ouvert de $\mathbb{R}^n$. La fonction $f : U \to \mathbb{R}^p$ est de classe C^1 sur U si et seulement si f et toutes ses dérivées partielles sont définies et continues sur U.*

Preuve : $\Longrightarrow$ Supposons f de classe C^1. Alors on sait déjà que toutes ses dérivées partielles sont définies et l'on a de plus, en notant $(\epsilon_1, \cdots, \epsilon_n)$ la base canonique de $\mathbb{R}^n$,

$$\forall i \in \{1, \cdots, n\},\ \frac{\partial f}{\partial x_i}(a) = df(a)(\epsilon_i).$$

On en déduit que pour tout $(a, b) \in U^2$,

$$\frac{\partial f}{\partial x_i}(b) - \frac{\partial f}{\partial x_i}(a) = (df(b) - df(a))(\epsilon_i).$$

En revenant à la définition de $|||\cdot|||$, on a donc

$$\left\| \frac{\partial f}{\partial x_i}(b) - \frac{\partial f}{\partial x_i}(a) \right\|_F \leq |||df(b) - df(a)|||\ \|\epsilon_i\|_E,$$

ce qui assure la continuité de l'application $\frac{\partial f}{\partial x_i}$ sur U.
$\Longleftarrow$ D'après le théorème 3.3.1, f est différentiable sur U et l'on a

$$\forall x \in U,\ df(x) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) \circ \pi_i.$$

Comme π_i et $\frac{\partial f}{\partial x_i}$ sont des fonctions continues, le théorème de composition assure la continuité de df sur U. ∎

Proposition 3.3.3 *Soit U un ouvert de E, $a \in U$ et $f \in \mathcal{F}(U; F)$ une application continue sur U et différentiable sur $U \backslash \{a\}$. Si de plus $df(x)$ admet une limite $L \in \mathcal{L}(E; F)$ quand x tend vers a alors f est différentiable en a et $df(a) = L$.*

Preuve : Posons $\phi(x) = f(x) - L(x)$. Il suffit de démontrer que ϕ est différentiable en a et que $d\phi(a) = 0$.

Fixons $\epsilon > 0$. Par hypothèse, $d\phi(x)$ tend vers 0 quand x tend vers a. Il existe donc $\eta > 0$ tel que $|||d\phi(y)||| \leq \epsilon$ pour tout $y \in B(a, \eta)$.

En appliquant l'inégalité des accroissements finis, on trouve donc

$$\forall y \in B(a, \eta),\ \|\phi(y) - \phi(a)\|_F \leq \epsilon \|y - a\|_E,$$

et l'on conclut que ϕ est différentiable en a et que $d\phi(a) = 0$. ■

Il est bien connu que pour une fonction définie et dérivable sur un intervalle de $\mathbb{R}$ être constant est équivalent à la nullité de la dérivée. Nous souhaitons généraliser ce résultat aux fonctions de *plusieurs* variables. Pour ce faire, nous avons besoin d'introduire la notion de *connexité*[1].

Définition 3.3.4 *Soit $(a, b) \in E^2$. On dit que la partie Γ de E est une* **ligne polygonale** *d'extrémités a et b s'il existe un nombre fini de points $a_0, \cdots, a_{k+1}$ de E tels que*

$$a_0 = a,\ a_{k+1} = b \text{ et } \Gamma = [a_0, a_1] \cup \cdots \cup [a_k, a_{k+1}].$$

Définition 3.3.5 *Soit U un ouvert de E. On dit que que U est* **connexe** *si pour tout couple (a, b) de points de U il existe une ligne polygonale d'extrémités a et b contenue dans U.*

Exemple important : On dit qu'une partie A de $\mathbb{R}^n$ est **convexe** si pour tout couple (x, y) de points de A le segment fermé $[x, y]$ est inclus dans A.

Vu la définition de connexité, il est clair que tout ouvert convexe est connexe.

Théorème 3.3.6 *Soit U un ouvert de E et $f \in \mathcal{F}(U; F)$.*

(i) Si f est constante sur U alors f est différentiable sur U et de différentielle nulle en tout point.

(ii) Si de plus U est connexe et f est différentiable sur U de différentielle nulle alors f est constante sur U.

Preuve : De la définition de la différentielle, on déduit facilement que si f est constante alors $df = 0$.

Pour prouver *(ii)*, fixons deux points a et b de U. Il s'agit de montrer que $f(b) = f(a)$.

Comme U est connexe, il existe un nombre fini de points $a_0, \cdots, a_{k+1}$ de E tels que $a_0 = a$, $a_{k+1} = b$ et $[a_j, a_{j+1}] \subset U$ pour tout $j \in \{0, \cdots, k\}$. La différentielle de f est nulle sur chaque segment $[a_j, a_{j+1}]$. On peut donc appliquer l'inégalité des accroissements finis avec $M = 0$ et conclure que $f(a_{j+1}) = f(a_j)$ pour tout $j \in \{0, \cdots, k\}$. Donc $f(b) = f(a)$. ■

Pour terminer ce chapitre, donnons une application **hors-programme** de l'inégalité des accroissements finis à la différentiation des suites de fonctions.

Théorème 3.3.7 *Soit U un ouvert convexe de E et a un point de U. Considérons une suite $(f_n)_{n\in\mathbb{N}}$ d'applications de U dans F vérifiant :*

(i) $(f_n(a))_{n\in\mathbb{N}}$ converge dans F,

(ii) chaque f_n est différentiable sur U,

(iii) $(df_n)_{n\in\mathbb{N}}$ converge uniformément sur tout borné de U vers une fonction $g : U \to \mathcal{L}(E; F)$.

Alors il existe une fonction f différentiable sur U de différentielle g et telle que $(f_n)_{n\in\mathbb{N}}$ converge simplement vers f sur U, et uniformément sur toute partie bornée de U.

[1] Dans ce cours, nous nous limiterons à la définition de la connexité pour les *ouverts*. Pour la définition dans le cas général, on pourra se reporter à [3].

Preuve : **1ère étape :** Convergence de la suite vers une fonction f.

Soit A une partie bornée de E. On peut sans nuire à la généralité se limiter au cas où A est l'intersection de U avec une boule ouverte de rayon $M > 0$ et de centre a (si bien que A est convexe).

En vertu de (*iii*), la suite $(df_n)_{n\in\mathbb{N}}$ vérifie le critère de Cauchy uniforme sur A. Donc il existe un rang $N \in \mathbb{N}$ tel que pour $q \geq N$ et $r \geq N$, on ait

$$\forall x \in A,\ |||df_r(x) - df_q(x)||| \leq \frac{\epsilon}{2M}.$$

Fixons $(x, x_0) \in A^2$. En appliquant l'inégalité des accroissements finis à la fonction $f_r - f_q$ entre x et x_0, on obtient

$$\|f_r(x) - f_q(x) - f_r(x_0) + f_q(x_0)\|_F \leq \frac{\epsilon}{3M}\|x - x_0\|_E. \tag{3.3}$$

En choisissant $x_0 = a$, et en appliquant la deuxième inégalité triangulaire, on obtient donc, compte tenu de $\|x - a\|_E \leq M$,

$$\|f_r(x) - f_q(x)\|_F < \frac{\epsilon}{3} + \|f_r(a) - f_q(a)\|_F.$$

En utilisant maintenant l'hypothèse (*i*), on conclut que $(f_r(x))_{r\in\mathbb{N}}$ est une suite de Cauchy de $F^{\mathbb{N}}$ donc converge vers une limite $f(x)$.

En faisant tendre q vers l'infini dans l'inégalité ci-dessus, on constate de plus que la convergence est uniforme sur A. Comme chaque f_r est continue (car différentiable), nous en concluons que f est continue sur U.

2ème étape : Étude de la différentiabilité de f.

Fixons $x_0 \in U$ et posons $A = U \cap B(x_0, 1)$. on a pour tout $q \in \mathbb{N}$ et $x \in A$,

$$\begin{aligned}\|f(x) - f(x_0) - g(x_0)(x - x_0)\|_F \leq \|(f(x) - f(x_0)) - (f_q(x) - f_q(x_0))\|_F \\ +\|f_q(x) - f_q(x_0) - df_q(x_0)(x - x_0)\|_F + \|(df_q(x_0) - g(x_0))(x - x_0)\|_F.\end{aligned}$$

En faisant tendre r vers $+\infty$ dans (3.3), on sait qu'il existe $q_1 \in \mathbb{N}$ tel que pour $q \geq q_1$, le premier terme du membre de droite soit inférieur à $\frac{\epsilon}{3}\|x - x_0\|_E$ pour tout $x \in A$.

En vertu de l'hypothèse (*iii*), il existe $q_2 \in \mathbb{N}$ tel que le dernier terme soit inférieur à $\frac{\epsilon}{3}\|x - x_0\|_E$ pour tout $x \in A$ et $q \geq q_2$.

Posons $q = \max(q_1, q_2)$. La différentiabilité de f_q assure que pour $\eta > 0$ assez petit, le second terme est aussi inférieur à $\frac{\epsilon}{3}\|x - x_0\|_E$ pour tout $x \in B(x_0, \eta)$.

Finalement, on a donc

$$\forall x \in B(x_0, \eta),\ \|f(x) - f(x_0) - g(x_0)(x - x_0)\|_F \leq \epsilon\|x - x_0\|_E,$$

d'où la différentiabilité de f en x_0, avec $df(x_0) = g(x_0)$. ■

Remarques :

- Si l'on remplace l'hypothèse *différentiable* par C^1, la conclusion du théorème ci-dessus demeure avec C^1 au lieu de différentiable.
- Nous laissons au lecteur le soin d'établir un résultat analogue pour les séries de fonctions différentiables (ou C^1).

Chapitre 4

Différentielles d'ordre supérieur

4.1 Définitions

Nous avons vu qu'à toute fonction $f \in \mathcal{F}(U;F)$ différentiable sur U on pouvait associer sa différentielle df qui est une fonction de $\mathcal{F}(U;\mathcal{L}(E;F))$. Comme $\mathcal{L}(E;F)$ est lui-même un espace vectoriel normé de dimension finie, on peut s'interroger sur la continuité ou la différentiabilité de df en tant que fonction de U vers $\mathcal{L}(E;F)$ ce qui conduit naturellement à la définition suivante :

Définition 4.1.1 *Soit $f \in \mathcal{F}(U;F)$ et $a \in U$. On dit que f est* **deux fois différentiable** *en a si f est différentiable sur un voisinage de a et si df est elle-même différentiable en a.*

On dit que f est deux fois différentiable sur U si f est deux fois différentiable en tout point de U.

Remarques :

- La **différentielle seconde** $[d(df)](a)$ est donc une *application linéaire* de E dans $\mathcal{L}(E;F)$. Par conséquent, pour tout $h \in E$, $[d(df)](a)(h)$ est une application linéaire de E vers F. Il est d'usage de confondre $[d(df)](a)$ avec l'application *bilinéaire* de $E \times E$ dans F notée $d^2f(a)$ et définie par la formule :

$$\forall (h,k) \in E \times E,\ d^2f(a)(h,k) \stackrel{\text{déf}}{=} \big\{[d(df)](a)(h)\big\}(k).$$

- Si $x \mapsto d^2f(x)$ est une fonction continue de U vers $\mathcal{L}^2(E \times E;F)$, on dit que f est de classe C^2 .
- Si la différentielle seconde d'une fonction f est elle-même différentiable, on dit que f est trois fois différentiable. Il est d'usage d'identifier la différentielle troisième avec une application *trilinéaire* de $E \times E \times E$ vers F notée $d^3f(a)$.

  En suivant le même procédé, on peut définir la différentielle d'ordre k, $d^kf(a)$ qui est assimilée à une application de $\mathcal{L}^k(E^k;F)$. Si de plus $x \mapsto d^kf(x)$ est continue, on dit que f est de **classe** C^k.

Exemples : 1) Pour une fonction $f : \mathbb{R} \to F$ l'existence de la différentielle seconde en un point est équivalente à l'existence de la dérivée seconde. On a de plus la formule :

$$\forall (h,k) \in \mathbb{R}^2,\ d^2f(a)(h,k) = hkf''(a).$$

2) Soit B une application bilinéaire de $E \times E$ vers F. Dans le chapitre 2, on a vu que

$$\forall (x_1,x_2) \in E^2,\ \forall (h_1,h_2) \in E^2,\ dB(x_1,x_2)(h_1,h_2) = B(x_1,h_2) + B(h_1,x_2).$$

Pour (h_1,h_2) fixé, la différentielle de l'application $\phi : (x_1,x_2) \mapsto B(x_1,h_2) + B(h_1,x_2)$ est

$$d\phi(x_1,x_2)(k_1,k_2) = B(k_1,h_2) + B(h_1,k_2).$$

On conclut que pour tout $x \in E \times E$,

$$d^2B(x)(h,k) = B(k_1,h_2) + B(h_1,k_2),$$

avec la notation évidente $h = (h_1,h_2)$ et $k = (k_1,k_2)$.

On retiendra que la différentielle seconde d'une application bilinéaire est constante.

Théorème 4.1.2 *Soit U un ouvert de E et $f \in \mathcal{F}(U;F)$ une application deux fois différentiable en $a \in U$. Alors $d^2f(a)$ est bilinéaire symétrique.*

Preuve : Il s'agit de montrer que $d^2f(a)(h,k) = d^2f(a)(k,h)$ pour tout $(h,k) \in E^2$. Pour cela, on définit la quantité

$$\Delta(h,k) = f(a+h+k) - f(a+h) - f(a+k) + f(a)$$

qui peut être interprétée comme une *différence finie d'ordre deux* autour de a et l'on cherche à établir que $\Delta(h,k) \sim d^2f(a)(h,k)$ pour $(h,k) \to (0,0)$. En utilisant le fait que $\Delta(h,k) = \Delta(k,h)$, nous obtiendrons alors $d^2f(a)(h,k) = d^2f(a)(k,h)$.

Fixons un $\epsilon > 0$. Par définition de la différentiabilité de df en a, il existe $\eta > 0$ tel que

$$\forall h \in B(0,2\eta),\ |||df(a+h) - df(a) - [d(df)](a)(h)||| \leq \epsilon \|h\|_E.$$

on a donc par définition de la norme triple,

$$(4.1) \qquad \forall h \in B(0,\eta), \forall k \in E\ \|df(a+h)(k) - df(a)(k) - d^2f(a)(h,k)\|_F \leq \epsilon\|h\|_E\|k\|_E.$$

Fixons $(h,k) \in B(0,\eta) \times B(0,\eta)$. Pour $t \in [0,1]$, posons

$$\psi(t) = f(a+h+tk) - f(a+tk) - td^2f(a)(h,k).$$

La fonction ψ est dérivable sur $[0,1]$ et un calcul facile montre que

$$\psi'(t) = \Big[(df(a+h+tk) - df(a))(k) - d^2f(a)(h+tk,k)\Big] - \Big[(df(a+tk) - df(a))(k) - d^2f(a)(tk,k)\Big].$$

En conséquence, d'après (4.1), on a

$$\forall t \in [0,1],\ \|\psi'(t)\|_F \leq \epsilon(\|h\|_E + 2\|k\|_E)\|k\|_E.$$

En appliquant le théorème des accroissements finis à ψ entre 0 et 1 et en remarquant que $\psi(1) - \psi(0) = \Delta(h,k) - d^2f(a)(h,k)$, on conclut donc que

$$\|\Delta(h,k) - d^2f(a)(h,k)\|_F \leq 2\epsilon(\|h\|_E + \|k\|_E)^2.$$

En échangeant les rôles de h et k et en utilisant la symétrie de Δ, on obtient également

$$\|\Delta(h,k) - d^2f(a)(k,h)\|_F \leq 2\epsilon(\|h\|_E + \|k\|_E)^2.$$

Donc, en vertu de la deuxième inégalité triangulaire,

$$\forall (h,k) \in B(0,\eta) \times B(0,\eta), \|d^2f(a)(h,k) - d^2f(a)(k,h)\|_F \leq 4\epsilon(\|h\|_E + \|k\|_E)^2.$$

Comme $d^2f(a)$ est bilinéaire, l'inégalité ci-dessus est en fait valable pour tout $(h,k) \in E^2$. Il ne reste plus qu'à faire tendre ϵ vers 0 pour conclure à l'égalité de $d^2f(a)(h,k)$ et de $d^2f(a)(k,h)$. ■

Remarque : Plus généralement, on peut montrer que si f est k fois différentiable en a alors $d^kf(a)$ est k-linéaire symétrique, c'est-à-dire que pour toute permutation σ de $\{1,\cdots,k\}$ et pour tout $(h_1,\cdots,h_k) \in E^k$, on a

$$d^kf(a)(h_{\sigma(1)},\cdots,h_{\sigma(k)}) = d^kf(a)(h_1,\cdots,h_k).$$

La preuve est du même type et repose sur l'utilisation de différences finies d'ordre k.

4.2 Dérivées partielles d'ordre supérieur

Dans cette section, nous nous limitons pour simplifier à l'étude d'une application f définie sur un ouvert U de $\mathbb{R}^n$ et à valeurs dans $\mathbb{R}$.

Nous avons déjà vu que f est de classe C^1 sur U si et seulement si toutes ses dérivées partielles $\frac{\partial f}{\partial x_i}$ sont des fonctions définies et continues sur U.

Si la j-ème application partielle associée à $\frac{\partial f}{\partial x_i}$ est elle-même dérivable en a, sa dérivée est notée $\frac{\partial^2 f}{\partial x_j \partial x_i}(a)$ (ou plus simplement $\frac{\partial^2 f}{\partial x_i^2}(a)$ si $i = j$). En appliquant la formule (2.4) à f puis à chaque dérivée partielle $\frac{\partial f}{\partial x_i}$, on obtient aisément le résultat suivant :

Proposition 4.2.1 *Soit $f : U \to \mathbb{R}$ différentiable sur U et deux fois différentiable en $a \in U$. Alors toutes les dérivées partielles d'ordre deux de f sont définies et l'on a*

$$\forall (h,k) \in \mathbb{R}^n \times \mathbb{R}^n,\ d^2 f(a)(h,k) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i k_j.$$

De la formule ci-dessus et de la symétrie de l'application bilinéaire $d^2 f(a)$, on déduit le résultat important suivant :

Théorème 4.2.2 (de Schwarz) *Soit U un ouvert de $\mathbb{R}^n$ et $f \in \mathcal{F}(U;\mathbb{R})$ une application deux fois différentiable en $a \in U$. Alors toutes les dérivées partielles d'ordre deux sont définies en a et l'on a*

$$\forall (i,j) \in \{1,\cdots,n\}^2,\ \frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

Définition 4.2.3 *Soit $f \in \mathcal{F}(U;\mathbb{R})$ une fonction dont toutes les dérivées partielles d'ordre deux sont définies en a. Alors la matrice*

$$D^2 f(a) \stackrel{\text{déf}}{=} \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(a) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) & \frac{\partial^2 f}{\partial x_2 \partial x_n}(a) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{pmatrix}$$

est appelée **matrice hessienne** *de f en a.*

Remarque 4.2.4 *En identifiant $h \in \mathbb{R}^n$ et $k \in \mathbb{R}^n$ aux vecteurs colonnes de mêmes composantes, on obtient la relation suivante entre $d^2 f(a)$ et $D^2 f(a)$:*

(4.2) $$\forall (h,k) \in \mathbb{R}^n \times \mathbb{R}^n,\ d^2 f(a)(h,k) = {}^T h D^2 f(a) k.$$

Autrement dit, $D^2 f(a)$ est la matrice de la forme bilinéaire $d^2 f(a)$ par rapport à la base canonique de $\mathbb{R}^n$. Le théorème de Schwarz assure de plus que la matrice hessienne est symétrique.

Si la j-ème application partielle associée à $\frac{\partial f}{\partial x_i}$ et la i-ème application partielle associée à $\frac{\partial f}{\partial x_j}$ sont simultanément dérivables en a, on peut légitimement se demander s'il y a toujours égalité entre $\frac{\partial^2 f}{\partial x_j \partial x_i}(a)$ et $\frac{\partial^2 f}{\partial x_i \partial x_j}(a)$. On vient de voir que cela est vrai lorsque $d^2 f(a)$ est définie.

En pratique malheureusement, il est souvent plus facile de calculer des dérivées partielles d'ordre deux que d'établir qu'une fonction est deux fois différentiable. **L'existence de toutes les dérivées partielles d'ordre deux en un point ne garantit pas la différentiabilité d'ordre deux en ce point**.

Le résultat suivant montre que la conclusion du théorème de Schwarz demeure sous des hypothèses – moins fortes que la différentiabilité d'ordre deux – portant uniquement sur les dérivées partielles. Pour simplifier, nous nous limitons au cas d'une fonction de *deux* variables.

Théorème 4.2.5 *Soit U un ouvert de $\mathbb{R}^2$, $(x_0,y_0)\in U$ et $f:U\to\mathbb{R}$. On suppose que $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ et $\frac{\partial^2 f}{\partial y\partial x}$ sont définies sur U et que $\frac{\partial^2 f}{\partial y\partial x}$ est continue en (x_0,y_0). Alors $\frac{\partial^2 f}{\partial x\partial y}(x_0,y_0)$ existe également et l'on a*

$$\frac{\partial^2 f}{\partial x\partial y}(x_0,y_0)=\frac{\partial^2 f}{\partial y\partial x}(x_0,y_0).$$

Preuve : Pour simplifier les notations, supposons que $(x_0,y_0)=(0,0)$. On pose $\ell=\frac{\partial^2 f}{\partial y\partial x}(0,0)$.

Soit $\eta>0$ tel que pour tout $(x,y)\in\mathbb{R}^2$ vérifiant $\max(|x|,|y|)\le\eta$, on ait

$$\left|\frac{\partial^2 f}{\partial y\partial x}(x,y)-\ell\right|\le\epsilon.$$

Pour $(h,k)\in[-\eta,\eta]^2$, posons

$$\Delta(h,k)=f(h,k)-f(h,0)-f(0,k)+f(0,0).$$

En appliquant l'égalité des accroissements finis à l'application $x\mapsto f(x,k)-f(x,0)$, on obtient $\theta\in]0,1[$ tel que

$$\Delta(h,k)=h\left(\frac{\partial f}{\partial x}(\theta h,k)-\frac{\partial f}{\partial x}(\theta h,0)\right).$$

Les hypothèses de l'énoncé assurent que la fonction $y\mapsto\frac{\partial f}{\partial x}(\theta h,y)$ est dérivable sur $[0,\eta]$. En appliquant à nouveau l'égalité des accroissements finis, on trouve donc un $\theta'\in]0,1]$ tel que

$$\Delta(h,k)=hk\frac{\partial^2 f}{\partial y\partial x}(\theta h,\theta' k).$$

On en déduit que pour tout couple (h,k) de réels non nuls tels que $\max(|h|,|k|)\le\eta$, on a

$$\left|\frac{\Delta(h,k)}{hk}-\ell\right|\le\epsilon.$$

Par ailleurs,

$$\frac{\Delta(h,k)}{hk}=\frac{1}{h}\left[\frac{f(h,k)-f(h,0)}{k}-\frac{f(0,k)-f(0,0)}{k}\right].$$

En faisant tendre k vers 0, on en déduit que

$$|h|<\eta\Longrightarrow\frac{1}{h}\left[\frac{\partial f}{\partial y}(h,0)-\frac{\partial f}{\partial y}(0,0)\right]\in[\ell-\epsilon,\ell+\epsilon].$$

Comme $\epsilon>0$ est arbitraire, on conclut que $\frac{\partial f}{\partial y}$ est dérivable par rapport à x en 0, de dérivée égale à ℓ. ■

Exercice : Montrer que la fonction f définie sur $\mathbb{R}^2$ par $f(x,y)=xy\left(\frac{x^2-y^2}{x^2+y^2}\right)$ si $(x,y)\ne(0,0)$ et $f(0,0)=0$ admet des dérivées partielles d'ordre 2 en tout point mais que

$$\frac{\partial^2 f}{\partial x\partial y}(0,0)\ne\frac{\partial^2 f}{\partial y\partial x}(0,0).$$

Pour les dérivées partielles d'ordre supérieur ou égal à deux, le lecteur retiendra le résultat suivant qui se démontre en appliquant le théorème de Schwarz de façon itérée et le corollaire 3.3.2 :

Théorème 4.2.6 *Une fonction $f : U \to \mathbb{R}$ est de classe C^k sur U si et seulement si toutes ses dérivées partielles d'ordre inférieur ou égal à k sont définies et continues sur U. on a alors pour toute permutation σ de $\{1, \cdots, k\}$,*

$$\frac{\partial^k f}{\partial x_{\alpha_1} \cdots \partial x_{\alpha_k}}(a) = \frac{\partial^k f}{\partial x_{\alpha_{\sigma(1)}} \cdots \partial x_{\alpha_{\sigma(k)}}(a)}.$$

Remarque : En raisonnant composante par composante, on peut facilement généraliser les résultats de cette section aux fonctions à valeurs *vectorielles*.

4.3 Formules de Taylor

4.3.1 Formule de Taylor avec reste intégral

Commençons par rappeler la formule de Taylor avec reste intégral pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$:

Théorème 4.3.1 *Soit I un intervalle de $\mathbb{R}$ et $(a,h) \in \mathbb{R}^2$ tel que a et $a+h$ soient dans I. Soit $p \in \mathbb{N}$ et $f \in \mathcal{C}^{p+1}(I;\mathbb{R})$. Alors on a la formule suivante :*

$$f(a+h) = f(a) + \sum_{k=1}^{p} \frac{h^k}{k!} f^{(k)}(a) + \frac{h^{p+1}}{p!} \int_0^1 (1-t)^p f^{(p+1)}(a+th)\, dt.$$

Preuve : Supposons que $h \neq 0$ (sinon la formule est triviale). Si f est de classe C^1, on a

$$f(a+h) - f(a) = \int_a^{a+h} f'(y)\, dy.$$

En posant $y = a + th$, on obtient donc

$$f(a+h) - f(a) = h \int_0^1 f'(a+th)\, dt, \tag{4.3}$$

qui est la formule voulue dans le cas $p = 0$.

Dans le cas général, on applique la formule d'intégration par parties itérée à la fonction 1 que l'on intègre successivement en $(t-1)$, $(t-1)^2/2$, etc, et à la fonction $t \mapsto f'(a+th)$. Il vient

$$\int_0^1 f'(a+th)\, dt = \left[\sum_{i=1}^{p} (-h)^{i-1} \frac{(t-1)^i}{i!} f^{(i)}(a+th) \right]_0^1 + \frac{(-1)^p h^p}{p!} \int_0^1 (t-1)^p f^{(p+1)}(a+th)\, dt.$$

Après quelques simplifications faciles, on obtient le résultat voulu. ■

Donnons maintenant la généralisation de la formule de Taylor avec reste intégral aux fonctions de plusieurs variables.

Théorème 4.3.2 *Soit U un ouvert de E et $(a,h) \in E^2$ tel que $[a, a+h] \subset U$. Soit $f \in \mathcal{F}(U;\mathbb{R})$ une fonction de classe C^{p+1}. Alors on a*

$$f(a+h) = f(a) + \sum_{k=1}^{p} \frac{d^k f(a)(h, \cdots, h)}{k!} + \int_0^1 \frac{(1-t)^p}{p!} d^{p+1} f(a+th)(h, \cdots, h)\, dt.$$

Preuve : Pour $t \in [0,1]$, on définit $\varphi(t) = f(a+th)$. Grâce au théorème de composition, on établit que φ est définie et de classe C^{p+1} sur $[0,1]$ et que l'on a

$$\begin{aligned}
&\varphi'(t) = df(a+th)(h),\\
&\varphi''(t) = d^2f(a+th)(h,h),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&\varphi^{(p+1)}(t) = d^{p+1}f(a+th)(h,\cdots,h).
\end{aligned}$$

Comme φ est une fonction de $[0,1]$ dans $\mathbb{R}$ de classe C^p, on peut lui appliquer la formule de Taylor avec reste intégral pour les fonctions d'une seule variable, entre 0 et 1. Cela donne

$$\varphi(1) = \sum_{k=0}^{p} \frac{\varphi^{(k)}(0)}{k!} + \int_0^1 \frac{(1-t)^p}{p!}\varphi^{(p+1)}(t)\,dt.$$

Il ne reste plus qu'à remplacer les $\varphi^{(j)}$ par leur valeur pour obtenir l'égalité souhaitée. ■

Remarque : En appliquant le résultat ci-dessus composante par composante, on peut généraliser la formule de Taylor d'ordre 2 avec reste intégral au cas d'une fonction à valeurs *vectorielles*.

Exemple : Pour une fonction C^2 de $\mathbb{R}^n$ dans $\mathbb{R}$, la formule de Taylor avec reste intégral se récrit :

$$f(a+h) = f(a) + \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a) + \sum_{i,j=1}^{n} h_i h_j \int_0^1 (1-t)\frac{\partial^2 f}{\partial x_i \partial x_j}(a+th)\,dt.$$

4.3.2 Formule de Taylor-Lagrange

La formule de Taylor-Lagrange est une généralisation de l'égalité des accroissements finis mettant en jeu les dérivées d'ordre inférieur ou égal à $p+1$.

Théorème 4.3.3 *Soit U un ouvert de E et $(a,h) \in E^2$ tel que $[a, a+h] \subset U$. Soit $f \in \mathcal{F}(U;\mathbb{R})$ une fonction $p+1$ fois différentiable sur U. Alors il existe $\theta \in]0,1[$ tel que*

$$f(a+h) = f(a) + \sum_{k=1}^{p} \frac{d^k f(a)(h,\cdots,h)}{k!} + \frac{d^{p+1}f(a+\theta h)(h,\cdots,h)}{(p+1)!}.$$

Preuve : On introduit à nouveau la fonction $\varphi : t \mapsto f(a+th)$ et on lui applique la formule de Taylor-Lagrange pour les fonctions numériques d'une variable réelle. ■

Exemple : Pour une fonction deux fois différentiable de $\mathbb{R}^n$ dans $\mathbb{R}$, la formule de Taylor-Lagrange d'ordre 2 se lit :

$$f(a+h) = f(a) + \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a) + \frac{1}{2}\sum_{i,j=1}^{n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a+\theta h) \quad \text{avec} \quad \theta \in]0,1[.$$

Corollaire 4.3.4 *Soit $f \in \mathcal{F}(U;\mathbb{R})$ une fonction $p+1$ fois différentiable sur U. Supposons que $[a,b] \subset U$ et que $d^{p+1}f$ soit bornée par un réel $M \geq 0$ sur $]a,b[$. Alors on a l'inégalité suivante :*

$$\left| f(b) - f(a) - \sum_{k=1}^{p} \frac{d^k f(a)(b-a,\cdots,b-a)}{k!} \right| \leq M \frac{\|b-a\|_E^{p+1}}{(p+1)!}.$$

Preuve : on applique la formule de Taylor-Lagrange à f et l'on utilise le fait que

$$|d^{p+1}f(a+\theta(b-a))(b-a,\cdots,b-a)| \leq M\|b-a\|_E^{p+1}.$$

■

Remarque : La formule de Taylor-Lagrange est fausse en général pour les fonctions à valeurs vectorielles. On remarquera en effet que cette formule à l'ordre 1 n'est autre que l'égalité des accroissements finis. En revanche, le corollaire ci-dessus peut se généraliser aux fonctions à valeurs vectorielles. La preuve s'inspire de celle de l'inégalité des accroissements finis.

4.3.3 Formule de Taylor-Young

Contrairement aux deux autres formules de Taylor, celle de Taylor-Young n'est pas exacte : c'est un développement asymptotique.

Théorème 4.3.5 *Soit U un ouvert de E et $a \in U$. Soit $f \in \mathcal{F}(U;F)$ une fonction de classe C^{p-1} sur U et p fois différentiable en a. Alors il existe un voisinage V de 0 et une fonction $\varepsilon : V \to F$ tendant vers 0 en 0, tels que*

$$\forall h \in V,\ f(a+h) = f(a) + \sum_{k=1}^{p} \frac{d^k f(a)(h,\cdots,h)}{k!} + \varepsilon(h)\|h\|_E^p.$$

Preuve : La preuve se fait par récurrence sur p.

Le cas $p=1$ résulte de la définition de la différentiabilité.

Soit $p \geq 2$. Supposons le résultat vrai pour $p-1$ et prouvons-le pour p. Soit donc $f \in \mathcal{F}(U;F)$ une fonction de classe C^{p-1} qui est p fois différentiable en a.

Fixons un h tel que $[a,h] \subset U$ et posons

$$g(t) = f(a+th) - \sum_{k=1}^{p} \frac{t^k d^k f(a)(h,\cdots,h)}{k!}.$$

Vues les hypothèses, la fonction g est dérivable sur $[0,1]$ et l'on a

$$\begin{aligned} g'(t) &= df(a+th)(h) - \sum_{k=1}^{p} \frac{t^{k-1}}{(k-1)!} d^k f(a)(h,\cdots,h), \\ &= \left(df(a+th) - \sum_{j=0}^{p-1} \frac{t^j}{j!} d^j(df)(a)(h,\cdots,h) \right)(h). \end{aligned}$$

En appliquant à df la formule de Taylor-Young d'ordre $p-1$, on obtient l'existence d'un voisinage V de 0 (que l'on peut toujours supposer convexe) et d'une fonction ε définie sur V tendant vers 0 en 0, tels que

$$\forall h' \in V,\ df(a+h') = df(a) + \sum_{j=1}^{p-1} \frac{d^j(df)(a)(h',\cdots,h')}{j!} + \varepsilon(h')\|h'\|_E^{p-1}.$$

Pour $h \in V$, on peut appliquer la formule ci-dessus à $h' = th$ et on obtient

$$\forall t \in [0,1],\ \|g'(t)\|_F \leq \varepsilon(th) t^{p-1} \|h\|_E^p.$$

Enfin, on remarque

$$g(1) - g(0) = f(a+h) - f(a) - \sum_{k=1}^{p} \frac{d^k f(a)(h,\cdots,h)}{k!}$$

et il ne reste plus qu'à appliquer l'inégalité des accroissements finis à g pour conclure. ■

Exemple : Pour une fonction C^1 de $\mathbb{R}^n$ dans $\mathbb{R}$, deux fois différentiable en a, la formule de Taylor-Young se récrit :

$$f(a+h) = f(a) + \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a) + \frac{1}{2} \sum_{i,j=1}^{n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a) + \varepsilon(h)\|h\|_E^2.$$


exosup.com

Chapitre 5

Problèmes d'extrema

Dans ce chapitre, U désigne un ouvert de l'e.v. E de dimension n et f, une fonction de U à valeurs dans $\mathbb{R}$.

5.1 Définitions

Définition 5.1.1 *Supposons f différentiable en $a \in U$. On dit que la fonction f admet un* **point critique** *en $a \in U$ si $df(a) = 0$.*

Définition 5.1.2 *Soit A une partie quelconque de E. On dit que $f : A \to \mathbb{R}$ admet un* **maximum local** *en $a \in A$ s'il existe un voisinage V de a tel que*

$$\forall x \in V \cap A,\ f(x) \leq f(a).$$

Lorsque l'inégalité ci-dessus est vraie pour tout $x \in A$, on dit que f admet un **maximum global** *en a.*

On dit que f admet un **minimum local** *en $a \in A$ s'il existe un voisinage V de a tel que*

$$\forall x \in V \cap A,\ f(x) \geq f(a).$$

Lorsque l'inégalité ci-dessus est vraie pour tout $x \in A$, on dit que f admet un **minimum global** *en a.*

Si les inégalités ci-dessus sont strictes pour $x \neq a$, on parle de minimum (ou de maximum) **strict**.

Remarque : Le mot **extremum** désigne un maximum ou un minimum.

Le théorème suivant assure que tout minimum ou maximum d'une fonction différentiable est un point critique.

Théorème 5.1.3 *Soit U un ouvert de E. Supposons f différentiable en $a \in U$. Si f admet un minimum ou un maximum local en a alors $df(a) = 0$.*

Preuve : Fixons $v \in E$ et considérons la fonction φ définie au voisinage de 0 par $\varphi(t) = f(a + tv)$. Alors φ est dérivable en 0 de dérivée $\varphi'(0) = df(a)(v)$. Par ailleurs φ admet un extremum en 0 donc $\varphi'(0) = 0$ et l'on conclut donc que $df(a)(v) = 0$. ■

Remarque : Réciproque fausse (même pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$ d'ailleurs).

Attention : Pour une fonction f définie et continue sur un *fermé* F, et différentiable à l'intérieur de F, le critère ci-dessus ne permet de détecter que les extrema *intérieurs* à F. Rien n'empêche f d'avoir un minimum ou un maximum sur la frontière ∂F. Le lecteur pourra considérer la fonction $f(x, y) = x^2 + y^2$ définie sur le disque unité fermé pour s'en convaincre...

5.2 Résultats liés à la compacité

Bien souvent, des arguments de compacité permettent de préciser la nature d'un point critique. Rappelons tout d'abord ce résultat classique du cours de topologie :

Théorème 5.2.1 *Soit K une partie compacte*[1] *de E et f une fonction continue de K dans $\mathbb{R}$. Alors f est bornée et atteint ses bornes. Plus précisément, il existe $a \in K$ et $b \in K$ tels que*

$$f(a) = \inf_{x\in K} f(x) = \min_{x\in K} f(x) \quad \textit{et} \quad f(b) = \sup_{x\in K} f(x) = \max_{x\in K} f(x).$$

Preuve : Il suffit de se souvenir que l'image d'un compact par une application continue est un compact et qu'un ensemble compact est borné et contient ses bornes inférieures et supérieures. ■

Pour une fonction f définie sur E tout entier, on peut avoir recours au résultat suivant :

Proposition 5.2.2 *Soit f une fonction continue de E vers $\mathbb{R}$ telle que*

$$\lim_{\|x\|_E \to +\infty} f(x) = +\infty.$$

Alors f est minorée et atteint son minimum.

Preuve : Il existe un $A > 0$ tel que $f(x) > f(0)$ pour $\|x\|_E > A$. La restriction g de f à la boule fermée $\overline{B}(0, A)$ est une application continue sur un compact. Donc elle est minorée et atteint son minimum en un point $a \in \overline{B}(0, A)$. on a donc $f(a) \le f(x)$ pour tout $x \in \overline{B}(0, A)$, et, en particulier, $f(a) \le f(0)$. Comme par définition de A, on a $f(x) > f(0)$ pour $\|x\|_E > A$, on peut maintenant conclure que f admet un minimum global en a. ■

Remarque : En changeant f en $-f$, on établit qu'une fonction f tendant vers $-\infty$ à l'infini est majorée et atteint son maximum.

5.3 Cas des fonctions deux fois différentiables

Dans le cas $E = \mathbb{R}^n$ (pour simplifier), l'étude de la matrice hessienne peut permettre de préciser la nature d'un point critique d'une fonction deux fois différentiable :

Théorème 5.3.1 *Soit $f : U \to \mathbb{R}$ une fonction différentiable sur l'ouvert U. Soit a un point critique de f. Supposons de plus que f soit deux fois différentiable en a.*

(i) Si $D^2f(a)$ est définie positive (resp. définie négative) alors f admet un minimum (resp. maximum) local strict en a.

(ii) Si f admet un minimum (resp. maximum) local en a alors $D^2f(a)$ est positive (resp. négative).

(iii) Si $D^2f(a)$ n'est ni positive ni négative alors f n'admet pas d'extremum en a. On dit que f admet un **point selle** *en a.*

Preuve : Munissons $\mathbb{R}^n$ de la norme euclidienne canonique $\|\cdot\|$. On applique la formule de Taylor-Young d'ordre 2 en a. Sachant que $df(a) = 0$, il vient pour tout h dans un voisinage V de 0 :

$$f(a+h) - f(a) = \frac{1}{2}\sum_{i,j=1}^{n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a) + o(\|h\|^2). \tag{5.1}$$

[1] C'est-à-dire fermée bornée puisque l'on travaille en dimension finie.

La somme du terme de droite (sans le facteur $1/2$) est l'image du couple (h, h) par la forme quadratique de matrice $D^2 f(a)$. Si l'on suppose cette matrice définie positive, il existe donc un $c > 0$ tel que

$$\forall h \in \mathbb{R}^n, \ \frac{1}{2} \sum_{i,j=1}^{n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \geq 2c\|h\|^2.$$

De plus il existe $\eta > 0$ tel que pour tout $h \in \mathbb{R}^n$ vérifiant $\|h\| < \eta$, le terme $o(\|h\|^2)$ puisse être majoré par $c\|h\|^2$. On conclut que

$$\forall h \in B(0, \eta), \ f(a+h) - f(a) \geq c\|h\|^2$$

et donc f admet un minimum local strict en a.

Réciproquement, si f admet un minimum local en a, le membre de droite de (5.1) doit être positif. Par définition du $o(\|h\|^2)$, on en déduit que pour tout $\epsilon > 0$, il existe $\eta > 0$ tel que

$$\forall h \in B(0, \eta), \ \sum_{i,j=1}^{n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \geq -\epsilon\|h\|^2.$$

Par homogénéité, cette inégalité est en fait vraie pour tout $h \in \mathbb{R}^n$. En faisant tendre ϵ vers 0, on conclut à la positivité de la matrice hessienne.

En changeant f en $-f$, on obtient le résultat relatif au cas où f admet un maximum en a.

Pour prouver le dernier point du théorème, on remarque d'abord que la matrice $D^2 f(a)$ admet une valeur propre $\lambda^- < 0$ et une valeur propre $\lambda^+ > 0$. Notons h^- (resp. h^+) un vecteur propre de norme 1 pour la valeur propre λ^- (resp. λ^+). Donnons-nous également un $\epsilon > 0$ tel que $\epsilon < \min(|\lambda^-|, \lambda^+)$. La formule de Taylor d'ordre 2 en a nous donne un $\eta > 0$ tel que

$$\forall h \in B(0, \eta), \left| f(a+h) - f(a) - \frac{{}^T h D^2 f(a) h}{2} \right| \leq \frac{1}{2}\epsilon\|h\|^2.$$

En appliquant cette inégalité à $h = th^-$ avec $|t| < \eta$, on obtient

$$\left| f(a + th^-) - f(a) - \frac{t^2}{2}\lambda^- \right| \leq \epsilon \frac{t^2}{2}.$$

d'où $f(a + th^-) - f(a) < 0$ pour $|t| < \eta$ non nul.

Un calcul similaire montre que $f(a + th^+) - f(a) > 0$ pour $|t| < \eta$ non nul. ∎

Remarque : Si $D^2 f(a)$ est positive ou négative mais non inversible, on ne peut pas conclure en général. Pour s'en convaincre, on pourra étudier la nature du point critique 0 pour les fonctions $f : (x, y) \mapsto x^2 + x^4 y$ et $g : (x, y) \mapsto x^2 + x^4 y^2$.

Exemple : étude des points critiques d'une fonction de deux variables

Théorème 5.3.2 *Soit U un ouvert de $\mathbb{R}^2$, $f : U \to \mathbb{R}$ différentiable sur U et $a \in U$ un point critique de f. Supposons f deux fois différentiable en a et notons $\Delta(a)$ le* **hessien** *de f en a, c'est-à-dire le déterminant de la matrice hessienne de f en a. Alors on a les résultats suivants :*

(i) Si $\Delta(a) > 0$, et $\frac{\partial^2 f}{\partial x^2}(a) > 0$ ou $\frac{\partial^2 f}{\partial y^2}(a) > 0$ alors f admet un minimum local strict en a.

(ii) Si $\Delta(a) > 0$, et $\frac{\partial^2 f}{\partial x^2}(a) < 0$ ou $\frac{\partial^2 f}{\partial y^2}(a) < 0$ alors f admet un maximum local strict en a.

(iii) Si $\Delta(a) < 0$ alors f admet un point selle en a.

Preuve : Posons $p \stackrel{\text{déf}}{=} \frac{\partial^2 f}{\partial x^2}(a)$, $q \stackrel{\text{déf}}{=} \frac{\partial^2 f}{\partial x \partial y}(a)$ et $r \stackrel{\text{déf}}{=} \frac{\partial^2 f}{\partial y^2}(a)$ de telle sorte que

$$D^2 f(a) = \begin{pmatrix} p & q \\ q & r \end{pmatrix}.$$

Le polynôme caractéristique de $D^2 f(a)$ est $X^2 - (p+r)X + pr - q^2$. Ses deux racines λ_- et λ_+ sont réelles (car la matrice correspondante est symétrique) et vérifient

$$\lambda_- + \lambda_+ = p + r \quad \text{et} \quad \lambda_- \lambda_+ = pr - q^2.$$

Si l'on suppose que $pr - q^2 > 0$ alors λ_- et λ_+ sont de même signe. Si de plus $p > 0$ (cas *(i)*) alors on doit avoir $r > 0$ (puisque $pr > q^2$) et l'on a donc $\lambda_- > 0$ et $\lambda_+ > 0$ et $D^2 f(a)$ est donc définie positive. D'après le théorème 5.3.1 f admet donc un minimum local strict en a.
Si $p < 0$ (cas *(ii)*) alors on doit avoir $r < 0$ et l'on a donc $\lambda_- < 0$ et $\lambda_+ < 0$. Donc $D^2 f(a)$ est définie négative. D'après le théorème 5.3.1 f admet donc un maximum local strict en a.
Si $pr - q^2 < 0$, les deux valeurs propres de $D^2 f(a)$ sont non nulles et de signe opposé. La fonction f a donc un point selle en a. ■

5.4 Conditions d'extrema dans le cas convexe

Définition 5.4.1 *Soit U une partie convexe de E. On dit que la fonction $f \in \mathcal{F}(U; E)$ est* **convexe** *si*

$$\forall (x, y) \in U^2, \, \forall t \in [0,1], \; f(tx + (1-t)y) \leq t f(x) + (1-t) f(y).$$

On dit que f est **strictement convexe** *si l'inégalité ci-dessus est stricte pour $x \neq y$ et $t \in]0,1[$.*

On dit que f est **concave** *(resp.* **strictement concave***) si $-f$ est convexe (resp. strictement convexe).*

5.4.1 Résultats de convexité pour les fonctions d'une seule variable

Le résultat suivant stipule que les pentes des segments reliant deux points du graphe d'une fonction convexe d'une seule variable réelle croissent avec les abscisses des points considérés. Cela constitue en fait une caractérisation des fonctions convexes d'une seule variable.

Lemme 5.4.2 (des trois cordes) *Soit f définie sur un intervalle I de $\mathbb{R}$ et à valeurs dans $\mathbb{R}$. Alors f est convexe si et seulement si pour tout triplet $(u, v, w) \in I^3$ tel que $u < v < w$ on a*

$$\frac{f(v) - f(u)}{v - u} \leq \frac{f(w) - f(u)}{w - u} \leq \frac{f(w) - f(v)}{w - v}. \tag{5.2}$$

De même, f est strictement convexe si et seulement si pour tout triplet $(u, v, w) \in I^3$ tel que $u < v < w$, les inégalités ci-dessus sont strictes.

Preuve : Supposons f convexe et donnons-nous $(u, v, w) \in I^3$ tel que $u < v < w$. Remarquons que $v = tu + (1-t)w$ avec $t = (w - v)/(w - u)$. Or par convexité de f, on a

$$f(v) \leq t f(u) + (1-t) f(w)$$

ce qui donne après multiplication par $w - u$,

$$(w - u) f(v) \leq (w - v) f(u) + (v - u) f(w).$$

On en déduit aisément les deux inégalités de (5.2).

Réciproquement, supposons que f vérifie (5.2) et donnons-nous $x < y$ et $t \in]0,1[$. En appliquant l'inégalité de gauche de (5.2) à $u = x$, $v = tx + (1-t)y$ et $w = y$, on obtient

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

Donc f est bien convexe.

Le cas strictement convexe se traite en remplaçant les inégalités par des inégalités strictes dans la preuve précédente. ■

Le résultat suivant assure que le graphe d'une fonction convexe dérivable se situe au-dessus de ses tangentes.

Proposition 5.4.3 *Soit $f : I \to \mathbb{R}$ dérivable en x_0.*

(i) Si f est convexe alors on a

$$\forall x \in I,\ f(x) \geq f(x_0) + (x - x_0)f'(x_0).$$

(ii) Si f est strictement convexe alors on a

$$\forall x \in I \setminus \{x_0\},\ f(x) > f(x_0) + (x - x_0)f'(x_0).$$

Preuve : Supposons d'abord que $x > x_0$. Soit $h \in]0, x - x_0[$. En appliquant l'inégalité de gauche du lemme des trois cordes à f avec $u = x_0$, $v = x_0 + h$ et $w = x$, on obtient

$$\frac{f(x_0 + h) - f(x_0)}{h} \leq \frac{f(x) - f(x_0)}{x - x_0}.$$

D'où l'inégalité voulue en faisant tendre h vers 0.

Si $x < x_0$, on applique l'inégalité de droite du lemme des trois cordes en $u = x$, $v = x_0 - h$ et $w = x_0$ avec $h \in]0, x_0 - x[$. Il vient

$$\frac{f(x_0) - f(x_0 - h)}{h} \geq \frac{f(x) - f(x_0)}{x - x_0}.$$

D'où l'inégalité voulue en faisant tendre h vers 0 (remarquer que multiplier par $x - x_0 < 0$ change le sens de l'inégalité).

Si f est strictement convexe et $x > x_0$, on commence par fixer un $x' \in]x_0, x[$. On sait que

$$f'(x_0) \leq \frac{f(x') - f(x_0)}{x' - x_0}.$$

Mais d'après le lemme des trois cordes pour les fonctions strictement convexes, on a

$$\frac{f(x') - f(x_0)}{x' - x_0} < \frac{f(x) - f(x_0)}{x - x_0},$$

d'où l'inégalité voulue. Le cas $x < x_0$ est similaire. ■

Donnons un corollaire fort utile du résultat précédent :

Corollaire 5.4.4 *Soit $f : I \to \mathbb{R}$ une fonction dérivable sur I. Alors f est convexe si et seulement si f' est une fonction croissante, et f est strictement convexe si et seulement si f' est strictement croissante.*

Preuve : Supposons f convexe. Soit $(x_0, x) \in I^2$ tel que $x_0 < x$. On applique la proposition précédente en x_0 puis en x et l'on obtient

$$f'(x_0) \leq \frac{f(x) - f(x_0)}{x - x_0} \leq f'(x).$$

Réciproquement, supposons f' croissante et donnons-nous un couple (x, y) tel que $x < y$ et $t \in]0, 1[$. On a

$$f(tx+(1-t)y) \leq tf(x)+(1-t)f(y) \iff \frac{f(x+(1-t)(y-x)) - f(x)}{1-t} \leq \frac{f(y) - f(y+t(x-y))}{t}.$$

En vertu de l'égalité des accroissements finis, il existe $c \in]x, x+(1-t)(y-x)[$ et $d \in]x+(1-t)(y-x), y[$ tels que

$$\frac{f(x + (1-t)(y-x)) - f(x)}{1-t} = (y-x)f'(c) \quad \text{et} \quad \frac{f(y) - f(y + t(x-y))}{t} = (y-x)f'(d).$$

Comme $c < d$, la croissance de f' assure donc que $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$. Le cas strictement convexe se traite en remplaçant les inégalités larges par des inégalités strictes. ■

On en déduit immédiatement le résultat suivant :

Corollaire 5.4.5 *Soit $f : I \to \mathbb{R}$ une fonction deux fois dérivable sur I. Alors f est convexe si et seulement si $f'' \geq 0$.*

Remarque : On ne peut pas caractériser la stricte convexité à l'aide de la dérivée seconde. On peut bien sûr affirmer que $f'' > 0$ entraîne la stricte convexité mais la réciproque est fausse (considérer $x \mapsto x^4$).

5.4.2 Résultats de convexité pour les fonctions de plusieurs variables

La propriété suivante des fonctions convexes se montre facilement par récurrence (exo : le faire).

Proposition 5.4.6 *Soit A un convexe de $\mathbb{R}^n$ et $f : A \to \mathbb{R}$ une fonction convexe. Alors pour tout $(x_1, \cdots, x_p) \in A^p$ et $(\alpha_1, \cdots, \alpha_p) \in [0, 1]^2$ vérifiant $\sum_{i=1}^{p} \alpha_i = 1$, on a*

$$f\left(\sum_{i=1}^{p} \alpha_i x_i\right) \leq \sum_{i=1}^{p} \alpha_i f(x_i).$$

Nous pouvons maintenant prouver le résultat suivant :

Théorème 5.4.7 *Soit f une fonction convexe définie sur un <u>ouvert</u> U de $\mathbb{R}^n$. Alors f est continue.*

Preuve : On munit $\mathbb{R}^n$ de la norme $\|x\| = \sum_{i=1} |x_i|$ (ce qui n'influe en rien sur les propriétés de continuité de f puisqu'en dimension finie toutes les normes sont équivalentes).

Fixons un $b \in U$. Il s'agit de montrer la continuité de f en b.

1. Réduction au cas $f(0) = 0$ et $\overline{B}(0, 1) \subset U$.

Quitte à considérer la fonction $g : x \mapsto f(\lambda x+b)-f(b)$ avec $\lambda > 0$ tel que $\overline{B}(b, \lambda) \subset U$, on peut se ramener à étudier la continuité en 0 pour une fonction convexe définie sur un ouvert contenant la boule unité fermée $\overline{B}(0, 1)$ et s'annulant en 0. On remarquera en effet que f et g sont simultanément convexes et que f est continue en b si et seulement si g est continue en 0.

2. Majoration de f sur $\overline{B}(0,1)$.

Notons a_0 l'origine et $a_i^{\pm} = (0,\cdots,0,\underbrace{\pm 1}_{i},0\cdots,0)$. On remarque que tout point $x=(x_1,\cdots,x_n)$ de $\overline{B}(0,1)$ se décompose en

$$x = (1-\|x\|)a_0 + \sum_{i=1}^{n} |x_i| a_i^{\epsilon_i} \ \text{ avec } \ \epsilon_i = \text{“+”} \ \text{ si } \ x_i \geq 0 \ \text{ et } \ \epsilon_i = \text{“–”} \ \text{ sinon.}$$

Il est clair que $1-\|x\| \in [0,1]$, $|x_i| \in [0,1]$ pour $i \in \{1,\cdots,n\}$ et $1-\|x\| + \sum_{i=1}^{n} |x_i| = 1$. Par convexité de f et comme $f(a_0)=0$, on a donc, d'après la proposition précédente,

$$f(x) \leq M\|x\| \ \text{ avec } \ M = \max(f(a_1^+),\cdots,f(a_n^+),f(a_1^-),\cdots,f(a_n^-)). \tag{5.3}$$

3. Fin de la preuve.

En écrivant que $0 = \frac{x}{2} + \frac{(-x)}{2}$, on obtient

$$f(x) \geq -f(-x) \geq -M\|x\|.$$

On conclut que

$$\forall x \in \overline{B}(0,1),\ |f(x)| \leq M\|x\|.$$

Donc f est continue en 0.

■

Proposition 5.4.8 *Soit U un ouvert convexe de $\mathbb{R}^n$ et $f : U \to \mathbb{R}$ convexe et différentiable en a. Alors on a*

$$\forall x \in U,\ f(x) \geq f(a) + df(a)(x-a).$$

Si f est strictement convexe et différentiable en a alors

$$\forall x \in U,\ (x \neq a) \Rightarrow f(x) > f(a) + df(a)(x-a)$$

Preuve : Soit $x \in U$. L'ouvert U est convexe donc contient le segment $[a,x]$. Cela permet de définir la fonction

$$\varphi \ : \ \begin{cases} [0,1] & \longrightarrow \ \mathbb{R} \\ t & \longmapsto \ f(a+t(x-a)). \end{cases}$$

La fonction φ est convexe et dérivable en 0. On peut donc lui appliquer la proposition 5.4.3 et l'on obtient en particulier

$$\varphi(1) \geq \varphi(0) + \varphi'(0).$$

Or $\varphi(0) = f(a)$, $\varphi(1) = f(x)$ et $\varphi'(0) = df(a)(x-a)$. Donc on obtient l'inégalité voulue. Le cas strictement convexe se traite en remplaçant l'inégalité ci-dessus par une inégalité stricte. ■

Interprétation géométrique. si $f : \mathbb{R} \to \mathbb{R}$ est convexe alors le graphe de f se situe *au dessus* de ses tangentes. Si $f : \mathbb{R}^2 \to \mathbb{R}$ est convexe, le graphe de f (qui est une surface de $\mathbb{R}^3$) se situe au dessus de ses plans tangents, etc.

Pour une fonction *deux fois différentiable*, l'étude de la matrice hessienne peut donner des renseignements sur la convexité :

Théorème 5.4.9 *Soit U un ouvert convexe de $\mathbb{R}^n$ et $f : U \to \mathbb{R}$ une fonction deux fois différentiable sur U.*

(i) La fonction f est convexe si et seulement si D^2f est une matrice positive en tout point.

(ii) Si D^2f est une matrice définie positive en tout point de U alors f est strictement convexe.

(iii) f est convexe sur U si et seulement si

$$\forall (a,x)\in U^2,\ d^2f(a)(x-a,x-a)\geq 0. \tag{5.4}$$

(iv) Si f vérifie

$$\forall (a,x)\in U^2,\ (x\neq a)\Rightarrow d^2f(a)(x-a,x-a)>0$$

alors f est strictement convexe sur U.

Preuve : Supposons f convexe et deux fois différentiable. Soit $a\in U$ et $h\in\mathbb{R}^n$. Il existe $t_0>0$ tel que $]a-t_0h,a+t_0h[\subset U$. On pose alors $\varphi(t)=f(a+th)$ pour $t\in]-t_0,t_0[$. Il est immédiat que φ est convexe et deux fois dérivable. Donc on a $\varphi''\geq 0$.

De plus, pour tout $t\in]-t_0,t_0[$, on a $\varphi'(t)=df(a+th)(h)$. Donc

$$\forall t\in]-t_0,t_0[,\ \varphi''(t)=d^2f(a+th)(h,h), \tag{5.5}$$

d'où $d^2f(a)(h,h)\geq 0$.

Comme h était arbitraire, on conclut que la forme bilinéaire symétrique $d^2f(a)$ est positive (ce qui est équivalent au fait que la matrice hessienne $D^2f(a)$ soit positive).

Réciproquement, supposons que $D^2f(y)$ soit une matrice positive en tout point y de U. Soit $(a,x)\in U^2$ tel que $x\neq a$. En posant $\varphi(t)=f(a+t(x-a))$ et en reprenant le calcul menant à (5.5), on obtient $\varphi''\geq 0$. Donc φ est convexe. Cela assure en particulier que

$$\forall t\in]0,1[,\ \varphi(t)\leq (1-t)\varphi(0)+t\varphi(1)$$

d'où

$$\forall t\in]0,1[,\ f(a+t(x-a))\leq (1-t)f(a)+tf(x).$$

Pour prouver (ii), il suffit de remplacer les inégalités par des inégalités strictes dans le calcul ci-dessus.

La partie directe de (iii) découle de (i). Pour prouver la réciproque, on utilise le fait que si f vérifie (5.4) alors pour tout $a\in U$, il existe $\eta>0$ tel que

$$\forall h\in\overline{B}(0,\eta),\ d^2f(a)(h,h)\geq 0.$$

En utilisant l'habituel raisonnement par homogénéité, on en déduit que l'inégalité ci-dessus est vraie pour tout $h\in\mathbb{R}^n$. Donc $d^2f(a)$ est positive. D'après (i), f est donc convexe.

La preuve de (iv) repose sur (ii). Les détails sont laissés au lecteur. ■

Remarque : En pratique, pour montrer qu'une fonction est convexe, on utilise souvent les faits évidents suivants :

- La somme de fonctions convexes est une fonction convexe.
- Une combinaison linéaire à coefficients positifs de fonctions convexes est convexe.
- Toute application linéaire ou affine à valeurs réelles est convexe.
- Toute forme quadratique positive est convexe.
- La convexité est préservée par changement de variable affine.

5.4.3 Extrema des fonctions convexes

Théorème 5.4.10 *Soit A un ensemble convexe (non nécessairement ouvert) non vide et f une fonction convexe sur A. Soit a un point intérieur à A. Supposons que f admette un point critique en a. Alors f admet un minimum global en a.*
Si f est strictement convexe, ce minimum est strict.

Preuve : Si A est ouvert, le résultat est une conséquence immédiate de la proposition 5.4.8. Étudions maintenant le cas général. Fixons une boule ouverte B non vide et centrée en a. Soit x un point de A. Pour t tendant vers 0, le point $a + t(x - a)$ tend vers a. Donc il existe un $t_0 \in]0, 1[$ tel que $a + t_0(x - a) \in B$. La fonction f restreinte à l'ouvert B est encore convexe et vérifie $df(a) = 0$. D'après la proposition 5.4.8, on a donc

$$f(a + t_0(x - a)) \geq f(a).$$

Mais par convexité de f, on a aussi

$$f(a + t_0(x - a)) \leq t_0 f(x) + (1 - t_0)f(a).$$

Donc $f(a) \leq t_0 f(x)+(1-t_0)f(a)$, ce qui entraîne visiblement $f(x) \geq f(a)$ puisque $t_0 > 0$. Si f est strictement convexe, toutes les inégalités ci-dessus sont strictes pour $x \neq a$ d'où la deuxième partie du théorème. ■

Remarque : Si f est concave, un résultat analogue est vrai en remplaçant minimum par maximum.

5.5 Exemple d'étude d'extrema

Soit $f : A \to \mathbb{R}$ une fonction. Indiquons comment procéder pour trouver les extrema de f.

1. Propriétés de compacité

Avant de se lancer dans le calcul de la matrice hessienne, il est prudent de vérifier si des arguments de compacité (du type théorème 5.2.1 ou proposition 5.2.2) ne permettent pas de conclure à l'existence d'au moins un extremum.

2. Propriétés de convexité

Si la fonction f est convexe, le théorème 5.4.10 assure l'existence (et éventuellement l'unicité) d'un extremum qui est en fait un minimum global.

3. Condition nécessaire d'extremum en un point intérieur à A

On calcule la jacobienne de f en tout point intérieur de A. Les points intérieurs où f admet un extremum sont à chercher parmi les points critiques, c'est-à-dire parmi les points $a \in \mathring{A}$ tels que $Df(a) = 0$.

4. Condition suffisante d'extremum en un point intérieur à A

Dans les cas "favorables", le calcul de $D^2 f$ combiné au théorème 5.3.1 permet de sélectionner les points critiques qui sont des extrema, et de préciser s'il s'agit de minima ou de maxima. Le calcul de la valeur de f en ces points et, éventuellement, l'étude du comportement de f à l'infini (si A n'est pas borné) permet de distinguer les extrema locaux des extrema globaux.

5. Points se situant sur la frontière de A

Si A n'est pas ouvert, il reste à examiner si f admet un extremum en un point situé sur la frontière de A. Il faut faire une étude au cas par cas. Des arguments de compacité peuvent éventuellement faciliter le raisonnement.

Exemple : Soit f la fonction définie sur $\mathbb{R}^2$ par $f(x,y) = x^4 + y^4 - 2(x-y)^2$. Cherchons à déterminer les extrema de f en suivant le schéma décrit ci-dessus.

1. Pour $\|(x,y)\|$ tendant vers l'infini, la fonction f se comporte comme $x^4 + y^4$ et tend donc vers l'infini. La proposition 5.2.2 assure donc que f admet (au moins) un minimum global.
2. La fonction f est la différence de deux fonctions convexes ce qui ne permet pas d'affirmer qu'elle est convexe (un calcul plus poussé montrerait qu'elle ne l'est pas).
3. On calcule
$$\frac{\partial f}{\partial x}(x,y) = 4x^3 - 4(x-y) \quad \text{et} \quad \frac{\partial f}{\partial y}(x,y) = 4y^3 - 4(y-x).$$
Donc $Df(x,y) = 0$ si et seulement si $x^3 = x - y = -y^3$. Donc (x,y) est un point critique si et seulement si
$$x = -y \quad \text{et} \quad x(x^2-2) = 0.$$
Donc f a trois points critiques :
$$(0,0), \quad (\sqrt{2}, -\sqrt{2}) \quad \text{et} \quad (-\sqrt{2}, \sqrt{2}).$$
4. On calcule
$$f(0,0) = 0, \quad f(\sqrt{2}, -\sqrt{2}) = -8 \quad \text{et} \quad f(-\sqrt{2}, \sqrt{2}) = -8.$$
Donc f atteint son minimum global -8 en deux points : $(\sqrt{2}, -\sqrt{2})$ et $(-\sqrt{2}, \sqrt{2})$. Reste à déterminer la nature du point $(0,0)$. Pour cela, on peut calculer la hessienne de f :
$$D^2 f(x,y) = \begin{pmatrix} 12x^2 - 4 & 4 \\ 4 & 12y^2 - 4 \end{pmatrix}.$$
Malheureusement le hessien de f en $(x,y) = (0,0)$ est nul, et le théorème 5.3.2 ne donne donc pas de renseignement sur la nature de $(0,0)$.
On peut éviter le calcul (inutile ici) de $D^2 f(0,0)$ en remarquant que $f(x,x) > 0$ pour $x \neq 0$ alors que $f(x,-x) < 0$ pour x assez petit, ce qui montre que $(0,0)$ n'est pas un extremum (il s'agit d'un point selle).

Chapitre 6

Fonctions implicites et inversion locale

6.1 Le théorème du point fixe

Dans cette partie, nous présentons deux théorèmes de point fixe contractant. Pour simplifier, nous nous limitons à l'énoncé dans le cas d'une fonction définie sur un fermé d'un espace vectoriel normé de dimension finie. Nous renvoyons le lecteur à [5] ou [7] pour une formulation plus générale.

Théorème 6.1.1 (du point fixe ou de Picard) *Soit Y une partie fermée d'un e.v.n. F de dimension finie et Φ une application k-contractante (c'est-à-dire k-lipschitzienne avec $k<1$) de Y dans Y. Alors Φ admet un unique point fixe.*

Preuve : L'unicité du point fixe est évidente.

Pour montrer l'existence, fixons un point x_0 de Y. On définit alors la suite $(x_n)_{n\in\mathbb{N}}$ par la relation de récurrence

$$\forall n\in\mathbb{N},\ x_{n+1}=\Phi(x_n).$$

Par une récurrence immédiate, on établit que

$$\forall n\in\mathbb{N},\ \|x_{n+1}-x_n\|_F\leq k^n\|x_1-x_0\|.$$

En conséquence, on a pour tout $n\in\mathbb{N}$ et $p\in\mathbb{N}^*$,

$$\|x_{n+p}-x_n\|_F\leq\frac{k^n}{1-k}\|x_1-x_0\|_F.$$

Donc $(x_n)_{n\in\mathbb{N}}$ est une suite de Cauchy. Comme F est un e.v.n complet car de dimension finie, cette suite converge vers une limite x. L'hypothèse Y fermé assure que $x\in Y$. Enfin, en passant à la limite dans la relation $x_{n+1}=\Phi(x_n)$, on conclut que $x=\Phi(x)$. ■

Pour démontrer le théorème des fonctions implicites, nous aurons besoin d'une version un peu plus sophistiquée du théorème du point fixe rendant compte de la dépendance continue par rapport à un paramètre.

Théorème 6.1.2 (du point fixe à paramètre) *Soit E et F deux e.v.n. de dimension finie. Soit A un sous-ensemble (quelconque) de E et Y une partie fermée de F. Soit $\Phi : A\times Y\to Y$ telle que pour tout $y\in Y$ l'application $\lambda\mapsto\Phi(\lambda,y)$ soit continue. Supposons de plus qu'il existe $k<1$ tel que pour tout $\lambda\in A$, $y\mapsto\Phi(\lambda,y)$ soit k-contractante.*

Alors pour tout $\lambda\in A$, il existe un unique $y_\lambda\in X$ vérifiant $\Phi(\lambda,y_\lambda)=y_\lambda$ et l'application $\lambda\mapsto y_\lambda$ est continue sur A.

Preuve : À λ fixé, l'application $y \mapsto \Phi(\lambda, y)$ vérifie les hypothèses du théorème du point fixe donc il existe un unique $y_\lambda \in Y$ vérifiant $\Phi(\lambda, y_\lambda) = y_\lambda$.

De plus, si $(\lambda, \mu) \in A^2$, on a (en notant $\|\cdot\|_E$ la norme sur E),

$$\begin{aligned} \|y_\lambda - y_\mu\|_E &= \|\Phi(\lambda, y_\lambda) - \Phi(\mu, y_\mu)\|_E, \\ &\leq \|\Phi(\lambda, y_\lambda) - \Phi(\lambda, y_\mu)\|_E + \|\Phi(\lambda, y_\mu) - \Phi(\mu, y_\mu)\|_E, \\ &\leq k\|y_\lambda - y_\mu\|_E + \|\Phi(\lambda, y_\mu) - \Phi(\mu, y_\mu)\|_E. \end{aligned}$$

De la dernière ligne, on déduit que

$$\|y_\lambda - y_\mu\|_E \leq \frac{1}{1-k}\|\Phi(\lambda, y_\mu) - \Phi(\mu, y_\mu)\|_E.$$

Or, à μ fixé, l'application $\lambda \mapsto \Phi(\lambda, y_\mu)$ est continue. Donc on obtient

$$\lim_{\lambda \to \mu} \|y_\lambda - y_\mu\|_E = 0.$$

■

6.2 Le théorème des fonctions implicites

Notation : Dans toute cette partie, E et F sont des e.v.n de dimension finie, et Ω désigne un ouvert de $E \times F$. Pour $f : \Omega \to F$ différentiable et $(x, y) \in E \times F$, on note $d_y f(x, y)$ la différentielle partielle en (x, y) de f par rapport à la seconde variable, c'est-à-dire la différentielle de l'application

$$\widetilde{f_x} \ : \left\{ \begin{array}{lcl} \Omega_x & \longrightarrow & F \\ y & \longmapsto & f(x, y), \end{array} \right.$$

où $\Omega_x \stackrel{\text{déf}}{=} \{y \in F \mid (x, y) \in \Omega\}$. On notera que Ω_x est ouvert si bien que parler de différentiabilité de $\widetilde{f_x}$ a bien un sens.

Nous utiliserons également la notation $d_x f$ pour désigner la différentielle de f par rapport à la *première* variable.

Théorème 6.2.1 (des fonctions implicites) *Soit Ω un ouvert de $E \times F$ et $f : \Omega \to F$ une application de classe C^1.*

Supposons qu'il existe $(x_0, y_0) \in \Omega$ tel que $f(x_0, y_0) = 0$ et $d_y f(x_0, y_0)$ soit un endomorphisme inversible. Alors il existe un voisinage ouvert U de x_0, un voisinage ouvert V de y_0 et une fonction $\varphi : U \to V$ de classe C^1 tels que

$$\Big((x, y) \in U \times V \quad \text{et} \quad f(x, y) = 0\Big) \iff y = \varphi(x).$$

De plus,

$$\forall x \in U,\ d\varphi(x) = -\big(d_y f(x, \varphi(x))\big)^{-1} \circ d_x f(x, \varphi(x)). \tag{6.1}$$

Preuve : **1ère étape : Résolubilité locale de l'équation** $f(x, y) = 0$

Pour $(x, y) \in \Omega$, posons $g(x, y) = y - \big(d_y f(x_0, y_0)\big)^{-1} f(x, y)$.

Le théorème de composition assure que g est de classe C^1 sur Ω. De plus il est clair que

$$f(x, y) = 0 \iff g(x, y) = y.$$

Donc notre problème se ramène à la détermination des points fixes de g avec x jouant le rôle d'un paramètre. Pour cela, nous allons faire appel au théorème du point fixe à paramètre après avoir vérifié qu'il s'applique à g.

Tout d'abord, g est clairement continue par rapport à la première variable. Ensuite, un calcul facile montre que

$$\forall (x,y) \in \Omega,\ d_y g(x,y) = \mathrm{Id}_F - \big(d_y f(x_0,y_0)\big)^{-1} d_y f(x,y)$$

si bien que $d_y g(x_0,y_0) = 0$. Par continuité de $d_y g$, il existe donc $\alpha > 0$ tel que

$$\big(\|x - x_0\|_E \le \alpha \quad \text{et} \quad \|y - y_0\|_F \le \alpha\big) \Longrightarrow |||d_y g(x,y)||| \le \frac{1}{2}.$$

En vertu du théorème des accroissements finis, l'application g est donc 1/2-contractante par rapport à y sur le fermé $\overline{B}_E(x_0,\alpha) \times \overline{B}_F(y_0,\alpha)$.
Par ailleurs, pour $(x,y) \in \overline{B}_E(x_0,\alpha) \times \overline{B}_F(y_0,\alpha)$, on a

$$\begin{aligned}
\|g(x,y) - y_0\|_F &= \|g(x,y) - g(x_0,y_0)\|_F,\\
&\le \|g(x,y) - g(x,y_0)\|_F + \|g(x,y_0) - g(x_0,y_0)\|_F,\\
&\le \tfrac{1}{2}\|y - y_0\|_F + \|g(x,y_0) - g(x_0,y_0)\|_F,\\
&\le \tfrac{\alpha}{2} + \|g(x,y_0) - g(x_0,y_0)\|_F.
\end{aligned}$$

Par continuité de g, il existe $\alpha' > 0$ tel que le dernier terme soit majoré par $\alpha/2$ pour $x \in \overline{B}_E(x_0,\alpha')$. L'application g restreinte à $\overline{B}_E(x_0,\alpha') \times \overline{B}_F(y_0,\alpha)$ est donc à valeurs dans $\overline{B}_F(y_0,\alpha)$ et le théorème du point fixe à paramètre s'applique.
On en déduit qu'il existe une application $\widetilde{\varphi}$ définie et continue sur $\overline{B}_E(x_0,\alpha')$ à valeurs dans $\overline{B}_F(y_0,\alpha)$ et telle que

$$\forall (x,y) \in \overline{B}_E(x_0,\alpha') \times \overline{B}_F(y_0,\alpha),\ f(x,y) = 0 \iff y = \widetilde{\varphi}(x).$$

Posons $V = B_F(y_0,\alpha)$ et $U = \widetilde{\varphi}^{-1}(V)$. Comme $\widetilde{\varphi}$ est continue, l'ensemble U est un ouvert. Il ne reste plus qu'à choisir pour φ l'application $\widetilde{\varphi}$ restreinte à U et à valeurs dans V.

2ème étape : Différentiabilité de φ en x_0
Comme f est différentiable en (x_0,y_0), il existe une application ε définie sur un voisinage W de $(0,0)$ et tendant vers 0 en $(0,0)$ telle que

$$\forall (h,k) \in W,\ f(x_0+h, y_0+k) = f(x_0,y_0) + d_x f(x_0,y_0)(h) + d_y f(x_0,y_0)(k) + \|(h,k)\|_{E\times F}\, \varepsilon(h,k).$$

Choisissons $h = x - x_0$ et $k = \varphi(x) - \varphi(x_0)$. Comme φ est continue et vaut y_0 en x_0, il existe $\beta > 0$ et une application η définie sur $B_E(x_0,\beta)$ et tendant vers 0 en x_0 telle que pour tout $x \in B_E(x_0,\beta)$ on ait

$$\begin{aligned}
f(x,\varphi(x)) - f(x_0,y_0) = d_x f(x_0,y_0)(x - x_0)&\\
+ d_y f(x_0,y_0)(\varphi(x) - y_0) + \|(x - x_0, \varphi(x) - y_0)\|_{E\times F}\, \eta(x)&.
\end{aligned}$$

Par définition de φ, le membre de gauche est nul. Donc en appliquant $d_y f(x_0,y_0)^{-1}$, on obtient

$$\begin{aligned}
\varphi(x) - \varphi(x_0) = -\big[\big(d_y f(x_0,y_0)\big)^{-1} \circ d_x f(x_0,y_0)\big](x - x_0)&\\
+ \|(x - x_0, \varphi(x) - \varphi(x_0))\|_{E\times F} \big(d_y f(x_0,y_0)\big)^{-1}(\eta(x))&.
\end{aligned} \tag{6.2}$$

En utilisant la continuité des différentes applications linéaires mises en jeu, on montre l'existence de deux réels $A, B \ge 0$ tels que la norme du membre de droite soit majorée par

$$(A + B\|\eta(x)\|_F)\|x - x_0\|_E + B\|\eta(x)\|_F \|\varphi(x) - \varphi(x_0)\|_F.$$

Quitte à diminuer β, on peut toujours supposer que $B\|\eta(x)\|_F \leq 1/2$ pour tout $x \in B_E(x_0,\beta)$. Par conséquent,

$$\forall x \in B_E(x_0,\beta),\ \|\varphi(x)-\varphi(x_0)\|_F \leq (2A+1)\|x-x_0\|_E.$$

On conclut que le dernier terme du membre de droite de (6.2) est un $o(\|x-x_0\|_E)$ puis que φ est différentiable en x_0 avec

$$d\varphi(x_0) = -\big(d_y f(x_0,y_0)\big)^{-1} \circ d_x f(x_0,y_0).$$

3ème étape : L'application φ est C^1
Comme $d_y f$ est continue et inversible en (x_0,y_0), l'application $(x,y) \mapsto \det d_y f(x,y)$ est continue et non nulle près de (x_0,y_0). Par conséquent, quitte à diminuer le domaine de définition U de φ, et V, on peut toujours supposer que $d_y f(x,y)$ est inversible pour tout $(x,y) \in U \times V$. En reprenant le raisonnement de l'étape précédente en x, on montre que φ est différentiable en x et que

$$d\varphi(x) = -\big(d_y f(x,\varphi(x))\big)^{-1} \circ d_x f(x,\varphi(x)).$$

Enfin, vu les hypothèses de continuité sur df, le membre de droite est continu, et l'on conclut que φ est C^1 sur U. ■

Remarque : Si l'on remplace C^1 par C^k dans le théorème ci-dessus, la fonction implicite φ est de classe C^k. Ses différentielles successives s'obtiennent en différentiant la formule (6.1).

Exemples :

1. La condition $df_y(x_0,y_0)$ inversible est une condition suffisante mais pas nécessaire pour la résolubilité locale de l'équation $f(x,y)=0$. Prenons le cas de la fonction $f:(x,y)\mapsto x-y^3$. Alors l'équation $f(x,y)=0$ a une unique solution pour tout $x\in\mathbb{R}$. Il s'agit de $y=\sqrt[3]{x}$. On remarquera toutefois que l'application $x\mapsto\sqrt[3]{x}$ n'est pas C^1 au voisinage de 0 car non dérivable en 0.
2. Pour une fonction $f:\mathbb{R}^2\to\mathbb{R}$ de deux variables, la condition $\partial_y f(x_0,y_0)\neq 0$ revient à dire que la *tangente* à la courbe de niveau $f(x,y)=f(x_0,y_0)$ en (x_0,y_0) n'est pas verticale.
3. Soit $f:(x,y)\mapsto x^2+y^2-1$ et (x_0,y_0) un point du cercle unité distinct de $(1,0)$ et $(-1,0)$. Le théorème des fonctions implicites assure la résolubilité locale de l'équation $f(x,y)=0$ près de (x_0,y_0). Un calcul explicite donne pour solution $\varphi(x)=\mathrm{sgn}(y_0)\sqrt{1-x^2}$. En revanche $\frac{\partial f}{\partial y}$ s'annule en $(1,0)$ et $(-1,0)$ et l'on constate effectivement que l'équation $x^2+y^2-1=0$ a *deux* branches de solutions solutions (à savoir $\pm\sqrt{1-x^2}$) près de $(1,0)$ et de $(-1,0)$.
4. La formule donnant $d\varphi$ en fonction de $d_x f$ et $d_y f$ se retrouve aisément en différentiant la relation
$$\forall x\in U,\ f(x,\varphi(x))=0.$$

6.3 Extrema sous contraintes

Dans le chapitre précédent, nous avons étudié les extrema d'une fonction f définie sur un *ouvert* de $\mathbb{R}^n$. Nous avons vu que l'étude de df et de d^2f permettait d'obtenir des conditions nécessaires et des conditions suffisantes d'extremum.

Pour un ensemble Γ qui n'est pas ouvert, peut-on encore donner des critères permettant de déterminer les extrema de la restriction $f_{|\Gamma}$ de la fonction f à l'ensemble Γ ?

Dans cette section, nous donnons des éléments de réponse lorsque Γ est du type[1] $g^{-1}(\{0\})$ avec g fonction de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$. On parle alors d'**extrema sous contraintes**.

6.3.1 Le cas d'une seule contrainte

En guise d'échauffement, nous étudions d'abord le cas où la fonction g est à valeurs réelles.

Géométriquement, $\Gamma \stackrel{\text{déf}}{=} g^{-1}(\{0\})$ est une hypersurface de $\mathbb{R}^n$. On dit alors que Γ est défini à l'aide *d'une seule contrainte* et que l'on cherche les extrema de f sous la contrainte $g = 0$.

Théorème 6.3.1 *Soit U un ouvert de $\mathbb{R}^n$, f et g deux fonctions de classe C^1 sur U. Soit $\Gamma = g^{-1}(\{0\})$. Supposons que $f_{|\Gamma}$ admette un extremum local en un point a de Γ tel que $dg(a) \neq 0$. Alors il existe un unique $\lambda \in \mathbb{R}$ tel que*

$$d(f + \lambda g)(a) = 0.$$

Preuve : Notons $a' = (a_1, \cdots, a_{n-1})$ de telle sorte que $a = (a', a_n)$. Pour $x \in \mathbb{R}^n$, on pose $x = (x', x_n)$. Comme $dg(a) \neq 0$, on peut, quitte à renuméroter les variables, supposer que $\frac{\partial g}{\partial x_n}(a) \neq 0$.

En vertu du théorème des fonctions implicites, il existe donc un voisinage V de a et une fonction φ de classe C^1 définie sur un voisinage V' de a' et telle que

$$x \in V \cap \Gamma \iff x' \in V' \quad \text{et} \quad x_n = \varphi(x').$$

Pour $x' \in V'$, posons $F(x') = f(x', \varphi(x'))$. Dire que $f_{|\Gamma}$ a un extremum local en a revient à dire que F a un extremum local en a'. Par composition, F est de classe C^1. Donc on doit avoir $dF(a') = 0$. Un calcul simple montre que cela est équivalent à

$$\forall i \in \{1, \cdots, n-1\}, \frac{\partial f}{\partial x_i}(a) + \frac{\partial f}{\partial x_n}(a)\frac{\partial \varphi}{\partial x_i}(a') = 0.$$

Par ailleurs, on a $g(x', \varphi(x')) = 0$ pour tout $x' \in V'$ donc

$$\forall i \in \{1, \cdots, n-1\}, \ \frac{\partial g}{\partial x_i}(a) + \frac{\partial g}{\partial x_n}(a)\frac{\partial \varphi}{\partial x_i}(a') = 0.$$

Donc on a bien $d(f + \lambda g)(a) = 0$ avec

$$\lambda = -\frac{\frac{\partial f}{\partial x_n}(a)}{\frac{\partial g}{\partial x_n}(a)}.$$

Supposons maintenant qu'il existe un autre réel μ tel que $d(f + \mu g)(a) = 0$. on a donc $(\lambda - \mu)dg(a) = 0$. Comme $dg(a) \neq 0$, on doit avoir $\lambda = \mu$. ■

Remarques : Le réel λ est appelé multiplicateur de Lagrange associé à la contrainte $g = 0$.

Exemple : Cherchons les extrema de la fonction $f : (x, y) \mapsto x + y$ restreinte à l'ensemble $\Gamma = \{(x, y) \in \mathbb{R}^2 \mid x^4 + y^4 = 1\}$.

Par définition, $\Gamma = g^{-1}(\{0\})$ avec $g(x, y) = x^4 + y^4 - 1$. Il est clair que dg ne s'annule jamais sur Γ. Le théorème précédent nous assure donc que si $f_{|\Gamma}$ admet un extremum en $(a, b) \in \Gamma$ alors il existe $\lambda \in \mathbb{R}$ tel que $d(f + \lambda g)(a, b) = 0$. Un calcul immédiat montre que cette équation est équivalente à

$$1 + 4\lambda a^3 = 0 \quad \text{et} \quad 1 + 4\lambda b^3 = 0.$$

[1]Sous certaines conditions de non dégénérescence, Γ est un objet de dimension $n-p$ appelé sous-variété. Nous renvoyons au cours de maîtrise ou à [2, 9, 10] pour plus de détails.

Le cas $\lambda = 0$ est visiblement exclu. On en déduit alors facilement que $a = b$. Sachant que $(a,b) \in \Gamma$, cela laisse pour seules possibilités $(a,b) = (-2^{-\frac{1}{4}}, -2^{-\frac{1}{4}})$ ou $(a,b) = (2^{-\frac{1}{4}}, 2^{-\frac{1}{4}})$.

Réciproquement, Γ étant compact et f continue, on sait que f admet au moins un minimum global et un maximum global sur Γ. En calculant $f(-2^{-\frac{1}{4}}, -2^{-\frac{1}{4}})$ et $f(-2^{-\frac{1}{4}}, -2^{-\frac{1}{4}})$, on peut conclure que $f_{|\Gamma}$ atteint son minimum en $(-2^{-\frac{1}{4}}, -2^{-\frac{1}{4}})$ et son maximum en $(a,b) = (2^{-\frac{1}{4}}, 2^{-\frac{1}{4}})$.

Exercice : Déterminer les extrema locaux de la fonction $f : (x,y,z) \mapsto xe^{y+z}$ restreinte à la sphère unité $\Gamma : x^2 + y^2 + z^2 = 1$.

6.3.2 Le cas de plusieurs contraintes

Nous considérons maintenant le cas où g est à valeurs dans $\mathbb{R}^p$. On note $(g_1, \cdots, g_p)$ les composantes de g.

Théorème 6.3.2 *Soit U un ouvert de $\mathbb{R}^n$, f et g deux fonctions de classe C^1 sur U avec f à valeurs dans $\mathbb{R}$ et g à valeurs dans $\mathbb{R}^p$. Notons $\Gamma = g^{-1}(\{0\})$.*
Supposons que $f_{|\Gamma}$ admette un extremum local en un point a de Γ tel que $dg(a)$ soit surjective. Alors il existe p réels $\lambda_1, \cdots, \lambda_p$ (appelés multiplicateurs de Lagrange) tels que

$$d(f + \sum_{i=1}^{p} \lambda_i g_i)(a) = 0.$$

De plus, le p-uplet $(\lambda_1, \cdots, \lambda_p)$ est unique.

Preuve : L'hypothèse de surjectivité de $dg(a)$ signifie que $\mathrm{rg}\,(dg(a)) = p$. En conséquence il existe un sous-espace vectoriel $F \subset \mathbb{R}^n$ de dimension p tel que $\mathbb{R}^n = \ker dg(a) \oplus F$ et que la restriction $dg(a)_F$ de $dg(a)$ à F soit un isomorphisme de F sur $\mathbb{R}^p$.

Quitte à faire un changement de variable[2], on peut toujours supposer que $\ker dg(a) = \mathbb{R}^{n-p} \times \{0\}$ et $F = \{0\} \times \mathbb{R}^p$. Pour $x \in \mathbb{R}^n$, on pose $x = (y,z)$ avec $y \in \mathbb{R}^{n-p}$ et $z \in \mathbb{R}^p$. De la même façon, on définit $a = (b,c)$ avec $b \in \mathbb{R}^{n-p}$ et $c \in \mathbb{R}^p$.

Comme g est C^1, $g(a) = 0$ et $d_z g(a)$ est inversible, le théorème des fonctions implicites s'applique et nous fournit une application φ définie et continue sur un voisinage ouvert U de b à valeurs dans un voisinage ouvert V de c et telle que

$$\forall (y,z) \in U \times V,\ (y,z) \in \Gamma \iff z = \varphi(y).$$

Les hypothèses montrent donc que l'application $y \mapsto f(y, \varphi(y))$ admet un extremum local en b. Par le théorème de composition, on a

$$d_y f(a) + d_z f(a) \circ d\varphi(b) = 0 \quad \text{et} \quad d_y g(a) + d_z g(a) \circ d\varphi(b) = 0,$$

d'où, puisque $d_z g(a)$ est inversible,

$$d_y f(a) - \big(d_z f(a) \circ (d_z g(a))^{-1}\big) \circ d_y g(a) = 0.$$

Remarquons que l'on a bien évidemment $d_z f(a) - \big(d_z f(a) \circ (d_z g(a))^{-1}\big) \circ d_z g(a) = 0$ donc finalement,

$$df(a) - \big(d_z f(a) \circ (d_z g(a))^{-1}\big) \circ dg(a) = 0.$$

[2] On peut considérer un isomorphisme $u : \mathbb{R}^n \to \mathbb{R}^n$ tel que $u(\ker dg(a)) = \mathbb{R}^{n-p} \times \{0\}$ et $u(F) = \{0\} \times \mathbb{R}^p$. On constate que $d(g \circ u^{-1})(a) = dg(a) \circ u^{-1}$ donc $d(g \circ u^{-1})(a)$ est un isomorphisme de $\{0\} \times \mathbb{R}^p$ sur $\mathbb{R}^p$. Par ailleurs, $f_{|\Gamma}$ a un extremum en $a \in \Gamma$ si et seulement si $(f \circ u^{-1})_{|u(\Gamma)}$ a un extremum en $u(a)$, et $u(\Gamma)$ coïncide avec $\{x \in \mathbb{R}^n,\ g \circ u^{-1}(x) = 0\}$.

L'application $-d_z f(a) \circ (d_z g(a))^{-1}$ est une forme linéaire sur $\mathbb{R}^p$. En vertu du théorème de représentation de Riesz, il existe donc un vecteur λ de $\mathbb{R}^p$ tel que

$$\forall h \in \mathbb{R}^p,\ -d_z f(a) \circ (d_z g(a))^{-1}(h) = (\lambda \mid h).$$

En notant $(\lambda_1, \cdots, \lambda_p)$ les composantes de λ et en constatant que pour $k \in \mathbb{R}^n$, $dg(a)(k)$ est le vecteur de $\mathbb{R}^p$ de composantes $(dg_1(a)(k), \cdots, dg_p(a)(k))$, on conclut que

$$df(a) + \sum_{i=1}^{p} \lambda_i dg_i(a) = 0.$$

Supposons maintenant qu'il existe un autre p-uplet $(\mu_1, \cdots, \mu_p)$ tel que

$$d(f + \sum_{i=1}^{p} \mu_i g_i)(a) = 0.$$

on a donc

$$\sum_{i=1}^{p} (\lambda_i - \mu_i) dg_i(a) = 0.$$

Par surjectivité de $dg(a)$, les formes linéaires $dg_1(a), \cdots, dg_p(a)$ sont linéairement indépendantes. Donc $\lambda_i = \mu_i$ pour tout $i \in \{1, \cdots, p\}$. ■

Exemple : Soit $f : (x, y, z) \mapsto x + z$ et $\Gamma = g^{-1}(0, 0)$ avec $g(x, y, z) = (x^2 + y^2 + z^2 - 1, z - y)$.

On vérifie facilement que $dg(x, y, z)$ est surjective pour tout $(x, y, z) \in \Gamma$. Une condition nécessaire d'extremum en $(x, y, z) \in \Gamma$ est donc l'existence de $(\lambda_1, \lambda_2) \in \mathbb{R}^2$ tel que $d(f + \lambda_1 g_1 + \lambda_2 g_2)(x, y, z) = 0$, ou encore

$$\begin{cases} 1 + 2\lambda_1 x = 0, \\ 2\lambda_1 y - \lambda_2 = 0, \\ 1 + 2\lambda_1 z + \lambda_2 = 0. \end{cases}$$

De la première égalité, on déduit que $\lambda_1 \neq 0$ et en sommant les deux autres égalités, on obtient $1 + 2\lambda_1(y + z) = 0$. Comme $1 + 2\lambda_1 x = 0$, on peut maintenant déduire que $x = y + z$. Sachant que sur Γ, on a $y = z$ et $x^2 + y^2 + z^2 = 1$, on conclut que f_Γ ne peut avoir d'extremum qu'en

$$\Big(-\frac{\sqrt{6}}{3}, -\frac{\sqrt{6}}{6}, -\frac{\sqrt{6}}{6}\Big) \quad \text{ou} \quad \Big(\frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{6}\Big).$$

Réciproquement, Γ étant compact et f continue, les deux points trouvés doivent correspondre à un minimum global et un maximal global. Un calcul rapide montre que $f_{|\Gamma}$ atteint son minimum (global) en $(-\frac{\sqrt{6}}{3}, -\frac{\sqrt{6}}{6}, -\frac{\sqrt{6}}{6})$ et son maximum (global) en $(\frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{6})$.

6.3.3 Le cas convexe

Dans le cas convexe, les conditions nécessaires d'extrema sous contraintes deviennent suffisantes. Plus précisément, on a le résultat suivant :

Théorème 6.3.3 *Soit U un ouvert convexe de $\mathbb{R}^n$ et g une application de U à valeurs dans $\mathbb{R}^p$. Soit $f : U \mapsto \mathbb{R}$. Notons $\Gamma \stackrel{\text{déf}}{=} g^{-1}(0)$. Supposons qu'il existe un point a de Γ et un p-uplet $(\lambda_1, \cdots, \lambda_p)$ de $\mathbb{R}^p$ tels que*

(i) Les fonctions f et g sont différentiables en a,

(ii) La fonction $f + \sum_{i=1}^{p} \lambda_i g_i$ est convexe,

(iii) $df(a) + \sum_{i=1}^{p} \lambda_i dg_i(a) = 0$.

Alors la fonction f restreinte à Γ admet un minimum global en a.

Preuve : Puisque la fonction $h \stackrel{\text{déf}}{=} f + \sum_{i=1}^{p} \lambda_i g_i$ est convexe sur l'ouvert U, le théorème 5.4.10 assure que h admet un minimum (global) en a. Donc a fortiori $h_{|\Gamma}$ admet un minimum en a. Mais clairement $h_{|\Gamma} = f_{|\Gamma}$ d'où le résultat. ∎

Remarques :

– Dans le cas convexe, les hypothèses sur g sont beaucoup moins fortes que dans le théorème 6.3.2. Cela tient au fait que le résultat *ne repose pas* sur le théorème des fonctions implicites.

– On remarquera que la condition $f + \sum_{i=1}^{p} \lambda_i g_i$ convexe est assurée dès que toutes les fonctions f, g_1, $\cdots$, g_p sont convexes et $\lambda_i \geq 0$ pour tout $i \in \{1, \cdots, p\}$.

Exemple : Soit Γ le cercle unité de $\mathbb{R}^2$ (défini par $\Gamma = g^{-1}(\{0\})$ avec $g(x,y) = x^2 + y^2 - 1$, et $f : (x,y) \mapsto ax + by$ où a et b sont deux réels donnés tels que $(a,b) \neq (0,0)$.

On vérifie que, pour $(x,y) \in \Gamma$, on a

$$d(f + \lambda g)(x,y) = 0 \quad \text{si et seulement si} \quad ay = bx \quad \text{et} \quad 2\lambda x = -a.$$

On obtient alors les deux points suivants :

$$(x_1, y_1) = \left(\frac{a}{\sqrt{a^2+b^2}}, \frac{b}{\sqrt{a^2+b^2}} \right) \text{ et } (x_2, y_2) = \left(\frac{-a}{\sqrt{a^2+b^2}}, \frac{-b}{\sqrt{a^2+b^2}} \right).$$

Le multiplicateur de Lagrange associé au point (x_2, y_2) est positif. Il est donc clair que $f + \lambda g$ est convexe, et le théorème précédent assure donc que $f_{|\Gamma}$ atteint son minimum global en (x_2, y_2).

Le multiplicateur de Lagrange associé au point (x_1, y_1) est *négatif.* on a donc cette fois-ci $f + \lambda g$ concave. En appliquant le théorème précédent à $-f$ et $-g$ on en déduit que $f_{|\Gamma}$ atteint son maximum global en (x_1, y_1).

6.4 Théorèmes d'inversion

Définition 6.4.1 *Soit $k \geq 1$, U un ouvert de E et V un ouvert de F. On dit que f est un C^k-difféomorphisme de U sur V si*

(i) f est bijective de U sur V,

(ii) f est de classe C^k sur U,

(iii) f^{-1} est de classe C^k sur V.

Remarque : Si f est un C^1 difféomorphisme de U sur V, on a

$$\forall x \in U,\ f^{-1}(f(x)) = x \quad \text{et} \quad \forall y \in V,\ f(f^{-1}(y)) = y.$$

En appliquant le théorème de composition, on a donc

$$\forall x \in U,\ df^{-1}(f(x)) \circ df(x) = \mathrm{Id}_E \quad \text{et} \quad \forall y \in V,\ df(f^{-1}(y)) \circ df^{-1}(y) = \mathrm{Id}_F.$$

Cela permet d'affirmer que $df(x)$ et $df^{-1}(f(x))$ sont inversibles. Autrement dit $df(x)$ est un isomorphisme [3] de E sur F et l'isomorphisme réciproque est $df^{-1}(f(x))$.

Le *théorème d'inversion locale* constitue une sorte de réciproque à ce résultat.

Théorème 6.4.2 (d'inversion locale) *Soit U un ouvert de E et $f : U \to E$ une application de classe C^k. Supposons qu'il existe $x_0 \in U$ tel que $df(x_0)$ soit inversible. Alors il existe un voisinage U' de x_0 et un voisinage V' de $y_0 \stackrel{\text{déf}}{=} f(x_0)$ tels que f soit un C^k-difféomorphisme de U' sur V'. De plus, en notant f^{-1} le difféomorphisme réciproque, on a*

$$\forall y \in V',\ df^{-1}(y) = \left[df(f^{-1}(y)) \right]^{-1}.$$

[3] Ce qui entraîne que E et F ont *même* dimension

Preuve : Définissons

$$\Phi : \left\{ \begin{array}{lcl} U \times E & \longrightarrow & E \\ (x,y) & \longmapsto & y - f(x). \end{array} \right.$$

L'application Φ est C^k sur $U \times E$ et l'on a $d_x\Phi(x,y) = -df(x)$. On constate donc que $\Phi(x_0, y_0) = 0$ et que $d_x\Phi(x_0,y_0)$ est inversible.

Le théorème des fonctions implicites donne l'existence d'un voisinage U' de x_0, d'un voisinage V' de y_0 et d'une fonction $\varphi \in C^k(V';U')$ tels que

$$\big((x,y) \in U' \times V' \quad \text{et} \quad y - f(x) = 0\big) \iff x = \varphi(y).$$

Comme $f(\varphi(y)) = y$, la fonction φ coïncide avec la bijection réciproque de f sur V'. De plus $df(\varphi(y))d\varphi(y) = \mathrm{Id}_E$ ce qui donne la formule voulue pour la différentielle de f^{-1}. ■

Remarque : Lorsque f est une fonction définie sur un ouvert de $\mathbb{R}^n$ et à valeurs dans $\mathbb{R}^n$, le calcul du jacobien donne une condition nécessaire et suffisante d'inversibilité pour la différentielle. Plus précisément,

$$df(x) \text{ est inversible} \iff \mathrm{Jac}_f(x) \neq 0.$$

Théorème 6.4.3 (d'inversion globale) *Soit U un ouvert de E et f une fonction de U dans E. On suppose que*

(i) f est de classe C^k sur U,

(ii) f est bijective de U sur $f(U)$,

(iii) Pour tout $x \in U$, $df(x)$ est inversible.

Alors $f(U)$ est un ouvert et f est un C^k-difféomorphisme de U sur $f(U)$.

Preuve : Soit $a \in U$. D'après le théorème d'inversion locale, il existe un ouvert U' contenant a et un ouvert V' contenant $f(a)$ tels que $f_{|U'}$ soit un difféomorphisme de U' sur V'. Cela assure que V' est un voisinage de $f(a)$ contenu dans $f(U)$. Donc $f(U)$ est un ouvert. L'hypothèse de bijectivité sur f nous permet de définir la bijection réciproque f^{-1} de f sur $f(U)$.

Clairement, le difféomorphisme réciproque de $f_{|U'}$ coïncide avec f^{-1} sur V'. Donc f^{-1} est de classe C^k. ■

Exemple : Le théorème d'inversion globale assure que la fonction $f : (r,\theta) \mapsto (r\cos\theta, r\sin\theta)$ est un difféomorphisme de $]0,+\infty[\times]-\pi,\pi[$ dans $\mathbb{R}^2 \setminus (]-\infty,0] \times \mathbb{R})$.

Pour les fonctions d'une seule variable et à valeurs réelles, le théorème d'inversion globale prend une forme particulièrement simple :

Corollaire 6.4.4 *Soit I un intervalle ouvert de $\mathbb{R}$ et $f : I \to \mathbb{R}$ une fonction de classe C^k telle que f' ne s'annule pas sur I. Alors $f(I)$ est un intervalle ouvert et f est un C^k-difféomorphisme de I sur $f(I)$.*

Preuve : Comme f' ne s'annule pas, l'égalité des accroissements finis assure que f est strictement monotone donc injective sur I. On peut donc appliquer le théorème d'inversion globale. Comme de plus l'image d'un intervalle de $\mathbb{R}$ par une application continue à valeurs réelles est un intervalle de $\mathbb{R}$, l'ensemble $f(I)$ est un intervalle ouvert. ■

6.5 Un peu de géométrie différentielle

6.5.1 Les hypersurfaces

Définition 6.5.1 *Soit Ω un ouvert de $\mathbb{R}^n$ et $f : \Omega \to \mathbb{R}$ une application de classe C^k. On dit que l'ensemble $\Sigma \stackrel{\text{déf}}{=} f^{-1}(0)$ est une* **hypersurface régulière** *de classe C^k de $\mathbb{R}^n$ si les deux conditions suivantes sont vérifiées :*

(i) Σ n'est pas vide,

(ii) pour tout $x \in \Sigma$, on a $df(x) \neq 0$.

On dit alors que $f(x) = 0$ est une **équation cartésienne** *de l'hypersurface Σ.*

Si $a \in \Sigma$, l'hyperplan affine de vecteur normal $\nabla f(a)$ et passant par a est appelé **hyperplan tangent** *à Σ en a.*

Exemple : Tout hyperplan de $\mathbb{R}^n$ est une hypersurface régulière.

Remarques : Soit Σ une hypersurface régulière de $\mathbb{R}^n$ d'équation $f(x) = 0$.

1. Si $n = 2$, on dit que Σ est une courbe régulière de $\mathbb{R}^2$.
2. Si $n = 3$, on dit que Σ est une surface régulière de $\mathbb{R}^3$.
3. Soit a un point de Σ. L'hyperplan H tangent à Σ en a a pour équation cartésienne

$$(x_1 - a_1)\frac{\partial f}{\partial x_1}(a) + \cdots + (x_n - a_n)\frac{\partial f}{\partial x_n}(a) = 0.$$

Le théorème suivant assure qu'une hypersurface régulière de $\mathbb{R}^n$ est, localement, un objet de dimension $n-1$ (i.e. peut être décrite au voisinage de chacun de ses points par $n-1$ paramètres).

Proposition 6.5.2 *Soit Σ une hypersurface régulière de classe C^k de $\mathbb{R}^n$, et $a \in \Sigma$. Alors il existe un voisinage ouvert $V \subset \mathbb{R}^n$ de a, un voisinage ouvert $U \subset \mathbb{R}^{n-1}$ de 0, et une application $\varphi : U \to V$ de classe C^k réalisant une bijection de U dans $V \cap \Sigma$, telle que $\varphi(0) = a$ et vérifiant* rg $d\varphi(h') = n - 1$ *pour tout $h' \in U$.*

Preuve : Soit $f(x) = 0$ une équation cartésienne de Σ au voisinage de a. Comme $df(a) \neq 0$, on peut, quitte à renuméroter les variables, supposer que $\frac{\partial f}{\partial x_n}(a) \neq 0$. Notons $a = (a', a_n)$ et $x = (x', x_n) = (a' + h', a_n + h_n)$. Pour h dans un voisinage de 0, on a

$$(a' + h', a_n + h_n) \in \Sigma \iff f(a' + h', a_n + h_n) = 0.$$

La n-ième dérivée partielle de l'application $h \mapsto f(a + h)$ en 0 vaut $\frac{\partial f}{\partial x_n}(a)$ donc n'est pas nulle. Le théorème des fonctions implicites nous donne donc deux ouverts U et V et une application ψ de classe C^k vérifiant

$$(a' + h', a_n + h_n) \in V \cap \Sigma \iff h' \in U \quad \text{et} \quad h_n = \psi(h').$$

Posons $\varphi_i(h') = a_i + h_i$ pour $i \in \{1, \cdots, n-1\}$, $\varphi_n(h') = a_n + \psi(h')$ et $\varphi = (\varphi_1, \cdots, \varphi_n)$. Par construction φ réalise une bijection de U dans $V \cap \Sigma$. De plus, les formes linéaires $(d\varphi_1(h'), \cdots, d\varphi_{n-1}(h'))$ sont indépendantes. Donc $d\varphi(h')$ est de rang $n - 1$. ■

Remarque : Tout point $x \in \Sigma \cap V$ peut donc s'exprimer ainsi :

$$\begin{cases} x_1 = \varphi_1(h_1, \cdots, h_{n-1}), \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x_n = \varphi_n(h_1, \cdots, h_{n-1}), \end{cases}$$

avec φ définie sur un voisinage U de 0 à valeurs dans un voisinage V de a, bijective de U vers $V \cap \Sigma$ et telle que rg $d\varphi(h') = n - 1$ pour $h' \in U$. On dit que $(h_1, \cdots, h_{n-1})$ constitue un système de **coordonnées locales** relatives au **paramétrage** φ.

Proposition 6.5.3 *Soit Σ une hypersurface régulière, $a \in \Sigma$ et φ un paramétrage de Σ au voisinage de a tel que $\varphi(0) = a$. Alors l'hyperplan tangent à Σ en a coïncide avec l'hyperplan affine H passant par a et dirigé par l'image de l'application linéaire $d\varphi(0)$. Autrement dit,*

$$H = a + \text{Vect}\left(\frac{\partial \varphi}{\partial h_1}(0), \cdots, \frac{\partial \varphi}{\partial h_{n-1}}(0)\right).$$

Preuve : Comme $f(x) = 0$ est une équation cartésienne de Σ qui est localement paramétrée par φ, on a $f \circ \varphi = 0$ près de 0. En conséquence,

$$df(a) \circ d\varphi(0) = 0,$$

d'où $\text{Im}\, d\varphi(0) \subset \text{Ker}\, df(a)$.

Mais $\text{Ker}\, df(a)$ est l'hyperplan orthogonal à $\nabla f(a)$ et $\text{Im}\, d\varphi(0)$ est de dimension $n-1$. Donc $\text{Im}\, d\varphi(0) = (\nabla f(a))^\perp$, d'où le résultat. ■

6.5.2 Application à la dimension 2

Rappelons tout d'abord la définition d'un arc paramétré de $\mathbb{R}^n$.

Définition 6.5.4 *On appelle* **arc paramétré** *de $\mathbb{R}^n$ toute application M continue sur un intervalle ouvert I de $\mathbb{R}$ et à valeurs dans $\mathbb{R}^n$. L'ensemble $\Gamma \stackrel{\text{déf}}{=} \{M(t) \mid t \in I\}$ est appelé support de l'arc paramétré. Si M est de classe C^k, on dit que l'arc paramétré est de classe C^k. On dit que l'arc est* **régulier** *si de plus M' ne s'annule pas sur I.*

Nous avons vu que toute application $\psi : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ de classe C^k et telle que $\nabla\psi$ ne s'annule pas sur $\psi^{-1}(\{0\})$ définissait une courbe Γ de $\mathbb{R}^2$ d'équation cartésienne $\psi(x, y) = 0$. D'après la proposition 6.5.2, il existe un voisinage V de (x_0, y_0) et deux applications X et Y définies sur un petit intervalle $]-\varepsilon, \varepsilon[$, telles que

(*i*) $X(0) = x_0,\ Y(0) = y_0$,

(*ii*) (X', Y') ne s'annule pas sur $]-\epsilon, \epsilon[$,

(*iii*) $\Sigma \cap V$ coïncide avec le support de l'arc paramétré $M(u) \stackrel{\text{déf}}{=} (X(u), Y(u))$ pour $u \in]-\epsilon, \epsilon[$.

La tangente en $(x_0, y_0) \in \Sigma$ est la droite d'équation cartésienne :

$$(x - x_0)\frac{\partial \psi}{\partial x}(x_0, y_0) + (y - y_0)\frac{\partial \psi}{\partial y}(x_0, y_0) = 0.$$

D'après la proposition 6.5.3, c'est aussi la droite passant par (x_0, y_0) et dirigée par $M'(0)$.

6.5.3 Application à la dimension 3

Donnons tout d'abord la définition d'une nappe paramétrée.

Définition 6.5.5 *On appelle* **nappe paramétrée** *de $\mathbb{R}^3$ toute application M continue d'un ouvert Ω de $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}^3$. Si M est de classe C^k, on dit que la nappe paramétrée est de classe C^k. L'ensemble $\Sigma \stackrel{\text{déf}}{=} \{M(u, v) \mid (u, v) \in \Omega\}$ est appelé support de la nappe paramétrée. On dit que la nappe est* **régulière** *si $dM(u, v)$ est de rang 2 pour tout $(u, v) \in \Omega$.*

Exemple : L'application M définie sur $\mathbb{R}^2$ par $M(r, \theta) = (\cos\theta\cos\phi, \cos\theta\sin\phi, \sin\theta)$ est une nappe paramétrée régulière. Son support est la sphère unité de $\mathbb{R}^3$.

Nous avons vu que toute application $\psi : \Omega \subset \mathbb{R}^3 \to \mathbb{R}$ de classe C^k et telle que $\nabla\psi$ ne s'annule pas sur $\psi^{-1}(\{0\})$ définissait une surface d'équation cartésienne $\psi(x, y, z) = 0$.

D'après la proposition 6.5.2, il existe un voisinage V de (x_0, y_0, z_0) et trois applications X, Y et Z définies sur un voisinage U de $(0, 0)$, telles que

(i) $X(0,0) = x_0,\ Y(0,0) = y_0,\ Z(0,0) = x_0$,

(ii) $\Sigma \cap V$ est le support de la nappe paramétrée $M(u,v) \stackrel{\text{déf}}{=} (X(u,v), Y(u,v), Z(u,v))$,

(iii) $\operatorname{rg} dM(u,v) = 2$ pour tout $(u,v) \in U$.

Le plan tangent P en $(x_0, y_0, z_0) \in \Sigma$ a pour équation cartésienne :

$$(x - x_0)\frac{\partial \psi}{\partial x}(x_0, y_0, z_0) + (y - y_0)\frac{\partial \psi}{\partial y}(x_0, y_0, z_0) + (z - z_0)\frac{\partial \psi}{\partial z}(x_0, y_0, z_0) = 0.$$

Si $(x_0, y_0, z_0) = M(u,v)$, ce plan tangent est aussi dirigé par les vecteurs $\frac{\partial M}{\partial u}(u,v)$ et $\frac{\partial M}{\partial v}(u,v)$.

Par conséquent $\frac{\partial M}{\partial u}(u,v) \wedge \frac{\partial M}{\partial v}(u,v)$ est un vecteur *normal* à P.

Exercice : Vérifier qu'une équation du plan tangent en $M(u,v)$ est donné par la relation

$$\begin{vmatrix} x - X(u,v) & \frac{\partial X}{\partial u}(u,v) & \frac{\partial X}{\partial v}(u,v) \\ y - Y(u,v) & \frac{\partial Y}{\partial u}(u,v) & \frac{\partial Y}{\partial v}(u,v) \\ z - Z(u,v) & \frac{\partial Z}{\partial u}(u,v) & \frac{\partial Z}{\partial v}(u,v) \end{vmatrix} = 0.$$

Pour terminer, nous allons étudier la nature géométrique de l'intersection de deux surfaces de $\mathbb{R}^3$. Nous allons établir que dans les cas “favorables”, cette intersection est localement le support d'un arc de $\mathbb{R}^3$.

Théorème 6.5.6 *Soit ψ_1 et ψ_2 deux applications C^k d'un ouvert Ω de $\mathbb{R}^3$ à valeurs réelles. Notons Σ_1 et Σ_2 les surfaces de $\mathbb{R}^3$ d'équations respectives $\psi_1(x,y,z) = 0$ et $\psi_2(x,y,z) = 0$. Supposons que $\Gamma \stackrel{\text{déf}}{=} \Sigma_1 \cap \Sigma_2$ ne soit pas vide et qu'il existe un point (x_0, y_0, z_0) de Γ tel que $\nabla\psi_1(x_0, y_0, z_0)$ et $\nabla\psi_2(x_0, y_0, z_0)$ ne soient pas colinéaires.*

Alors il existe un voisinage V de (x_0, y_0, z_0) tel que $\Gamma \cap V$ coïncide avec le support d'un arc de classe C^k et de vecteur tangent $\nabla\psi_1(x_0, y_0, z_0) \wedge \nabla\psi_2(x_0, y_0, z_0)$ en (x_0, y_0, z_0).

Preuve : Soit

$$\Psi \ : \ \begin{cases} \Omega \subset \mathbb{R}^3 & \longrightarrow \quad \mathbb{R}^2 \\ (x,y,z) & \longmapsto \quad (\psi_1(x,y,z), \psi_2(x,y,z)). \end{cases}$$

Par définition de Ψ, on a $\Gamma = \Psi^{-1}(\{0\})$ et $\operatorname{rg} d\Psi(x_0, y_0, z_0) = 2$. Autrement dit, la matrice jacobienne $D\Psi(x_0, y_0, z_0) \in \mathcal{M}_{2,3}(\mathbb{R})$ est de rang 2 donc contient une matrice extraite 2×2 inversible. Supposons pour fixer les idées que la matrice extraite constituée des colonnes $\frac{\partial \Psi}{\partial y}(x_0, y_0, z_0)$ et $\frac{\partial \Psi}{\partial z}(x_0, y_0, z_0)$ soit inversible. Le théorème des fonctions implicites fournit alors un $\eta > 0$, un voisinage V de (y_0, z_0) et une application $f = (Y, Z)$ de $]x_0 - \eta, x_0 + \eta[$ à valeurs dans V et de classe C^1 telle que $\Gamma \cap (]x_0 - \eta, x_0 + \eta[\times V)$ coïncide avec le support de l'arc paramétré $(x, Y(x), Z(x))$.

Dans le cas général, on peut par le même raisonnement paramétrer Γ près de (x_0, y_0, z_0) par l'une des trois coordonnées x, y ou z. Ainsi Γ coïncide avec le support d'un arc paramétré $M(u) = (X(u), Y(u), Z(u))$ tel que $M(u_0) = (x_0, y_0, z_0)$.

En dérivant par rapport à u les relations

$$\psi_i(X(u), Y(u), Z(u)) = 0$$

pour $i = 1, 2$, on obtient

$$X'(u)\frac{\partial \psi_i}{dx}(M(u)) + Y'(u)\frac{\partial \psi_i}{dy}(M(u)) + Z'(u)\frac{\partial \psi_i}{dz}(M(u)) = 0.$$

Autrement dit, $M'(u)$ est orthogonal aux vecteurs $\nabla\psi_1(u)$ et $\nabla\psi_2(u)$. Pour $u = u_0$, ces deux vecteurs sont indépendants donc $M'(u_0)$ est colinéaire à leur produit vectoriel. ■

Exemple : On retrouve que l'intersection de deux plans non parallèles est localement un arc paramétré. (En fait, dans ce cas, on sait bien évidemment que cette intersection est une droite !)

Chapitre 7

Introduction aux formes différentielles

Avant d'aborder l'étude des formes différentielles, il convient de faire quelques rappels sur les formes multi-linéaires alternées. Cette notion n'est pas tout à fait nouvelle puisqu'elle a déjà été introduite en deuxième année pour définir le déterminant. Dans la première partie de ce chapitre, nous donnons quelques résultats élémentaires d'algèbre extérieure. L'étude des formes différentielles sera faite dans la deuxième partie.

7.1 Quelques éléments d'algèbre extérieure

Dans toute cette partie, E désigne un $\mathbb{R}$-espace vectoriel de dimension finie n. On note E^* le **dual** de E, c'est-à-dire l'ensemble des formes linéaires sur E.

7.1.1 Définitions

À toute base $(e_1, \cdots, e_n)$ de E, on peut associer une base $(e_1^*, \cdots, e_n^*)$ de E^* appelée **base duale** de $(e_1, \cdots, e_n)$ et définie par

$$e_i^*(e_j) = 0 \quad \text{si} \quad i \neq j, \quad \text{et} \quad e_i^*(e_i) = 1.$$

Autrement dit, pour tout $x \in E$ s'écrivant $x = \sum_{j=1}^n x_j e_j$, on a $e_i^*(x) = x_i$.

Dans un souci de concision, on écrira désormais $e_i^*(e_j) = \delta_{ij}$ où δ_{ij} est le **symbole de Kronecker** défini par $\delta_{ij} = 0$ si $i \neq j$ et $\delta_{ii} = 1$.

Nous laissons au lecteur le soin de vérifier que $(e_1^*, \cdots, e_n^*)$ est bien une base de E^*.

Exemple : Dans $\mathbb{R}^n$, la base duale de la base canonique est la famille $(\pi_1, \cdots, \pi_n)$ constituée par les projections canoniques introduites dans le premier chapitre.

Rappelons maintenant la définition d'une **forme multi-linéaire alternée** :

Définition 7.1.1 *Soit $k \in \mathbb{N}^*$. On dit que φ est une* **forme k-linéaire alternée** *sur E si φ est une application de E^k vers $\mathbb{R}$ qui est k-linéaire (c'est-à-dire linéaire par rapport à chacune de ses variables) et vérifie*

$$\forall (v_1, \cdots, v_k) \in E^k, \ \varphi(v_1, \cdots, v_i, \cdots, v_j, \cdots, v_k) = -\varphi(v_1, \cdots, v_j, \cdots, v_i, \cdots, v_k)$$

pour tout couple d'entiers (i, j) tel que $1 \leq i < j \leq n$.

Exemples : 1) L'application qui à n vecteurs de E associe leur déterminant est une forme n-linéaire alternée sur E.

2) L'ensemble des formes bilinéaires alternées n'est autre que l'ensemble des formes bilinéaires antisymétriques.

Notation : On note $\Lambda_k(E)$ l'ensemble des formes k-linéaires alternées sur E.

La preuve du résultat suivant est laissée au lecteur :

Proposition 7.1.2 *Pour tout $k \in \mathbb{N}^*$, l'ensemble $\Lambda_k(E)$ est un espace vectoriel réel.*

Avant d'aller plus loin dans la théorie des formes k-linéaires alternées, nous avons besoin de rappeler ce que sont une permutation et la signature d'une permutation.

Définition 7.1.3 *On appelle* **permutation** *de n éléments toute application bijective de l'ensemble $\{1, \cdots, n\}$ dans lui-même. L'ensemble des permutations de n éléments est noté $\mathcal{S}_n$.*

Définition 7.1.4 *On dit que la permutation $\sigma \in \mathcal{S}_n$ est une* **transposition** *s'il existe deux entiers i et j tels que $1 \leq i < j \leq n$ et*

$$\sigma(k) = k \quad si \quad k \neq i \ et\ k \neq j, \quad \sigma(i) = j \quad et \quad \sigma(j) = i.$$

Rappelons que toute permutation est la composée d'un nombre fini de transpositions (voir par exemple [4]). La parité du nombre q de transpositions nécessaires pour décomposer une permutation σ donnée ne dépend pas de la décomposition choisie. On définit alors la **signature** ε_σ de σ par $\varepsilon_\sigma = (-1)^q$.

Proposition 7.1.5 *Soit $\phi \in \Lambda_k(E)$ et $(v_1, \cdots, v_k) \in E^k$. on a les propriétés suivantes :*

(i) Pour toute permutation σ de $\mathcal{S}_k$, on a $\varphi(v_{\sigma(1)}, \cdots, v_{\sigma(k)}) = \varepsilon_\sigma \varphi(v_1, \cdots, v_k)$.

(ii) Si $(v_1, \cdots, v_k) \in E^k$ est liée alors $\varphi(v_1, \cdots, v_k) = 0$.

Preuve : La première propriété se démontre en décomposant σ en produit de transpositions et en utilisant le caractère alterné. Il suffit de faire une récurrence sur le nombre de transpositions apparaissant dans la décomposition.

La deuxième propriété est évidente si la famille $(v_1, \cdots, v_k)$ contient deux fois le même vecteur (utiliser le caractère alterné). Le cas général découle alors de la linéarité par rapport à chaque variable. ■

Remarque : Supposons $k > n$. Alors toute famille de k vecteurs de E est liée. De la deuxième propriété de la proposition ci-dessus, on déduit donc que $\Lambda_k(E)$ est réduit à la fonction nulle.

7.1.2 Représentation des formes k-linéaires alternées

Dans cette partie, nous souhaitons obtenir une description un peu plus précise des éléments de $\Lambda_k(E)$. Pour cela, on fixe une fois pour toutes une base $(e_1, \cdots, e_n)$ de E.

- *Le cas* $k = 1$. L'ensemble $\Lambda_1(E)$ n'est autre que le dual E^* de E, donc est engendré par $(e_1^*, \cdots, e_n^*)$. Autrement dit, tout élément de $\Lambda_1(E)$ est de la forme

$$\varphi = \sum_{i=1}^{n} a_i e_i^* \quad \text{avec} \quad (a_1, \cdots, a_n) \in \mathbb{R}^n.$$

- *Le cas* $k = 2$. Soit φ une forme bilinéaire alternée, et $(x, y) \in E^2$. Décomposons x et y suivant la base $(e_1, \cdots, e_n)$:

$$x = \sum_{i=1}^{n} x_i e_i \quad \text{et} \quad y = \sum_{i=1}^{n} y_i e_i.$$

En vertu de la bilinéarité, puis du caractère alterné, on a

$$\begin{aligned}\varphi(x,y) &= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j \varphi(e_i,e_j),\\ &= \sum_{1\le i<j\le n} (x_i y_j - x_j y_i)\varphi(e_i,e_j).\end{aligned}$$

On constate que φ est entièrement déterminée par ses valeurs sur les couples (i,j) de $\{1,\cdots,n\}^2$ tels que $i<j$. En posant $\varphi_{ij} \stackrel{\text{déf}}{=} \varphi(e_i,e_j)$ et en notant $e_i^*\wedge e_j^*$ la forme bilinéaire alternée définie par

$$\forall (x,y)\in E^2,\ e_i^*\wedge e_j^*(x,y) = x_i y_j - x_j y_i,$$

on obtient donc

$$\varphi = \sum_{1\le i<j\le n} \varphi_{ij}\, e_i^*\wedge e_j^*.$$

On montre facilement (exercice : le faire) que la famille des $e_i^*\wedge e_j^*$ avec $1\le i<j\le n$ est linéairement indépendante. On en déduit que $\Lambda_2(E)$ est un espace vectoriel de dimension $n(n-1)/2$ et que l'ensemble des $e_i^*\wedge e_j^*$ avec $1\le i<j\le n$ en est une base.

- *Le cas général.* Pour tout $(i_1,\cdots,i_k)\in\{1,\cdots,n\}^k$, on définit la forme k-linéaire alternée $e_{i_1}^*\wedge\cdots\wedge e_{i_k}^*$ par

$$\forall (v_1,\cdots,v_k)\in E^k,\ e_{i_1}^*\wedge\cdots\wedge e_{i_k}^*(v_1,\cdots,v_k) = \sum_{\sigma\in\mathcal{S}_k} \varepsilon_\sigma e_{i_1}^*(v_{\sigma(1)})\cdots e_{i_k}^*(v_{\sigma(k)}).$$

Soit maintenant $\varphi\in\Lambda_k(E)$. Si l'on décompose chaque vecteur v_j en $v_j=\sum_{i=1}^n v_j^i e_i$, on obtient en vertu du caractère k-linéaire alterné de φ :

$$\varphi(v_1,\cdots,v_k) = \sum_{1\le i_1<\cdots<i_k\le n} \varphi(e_{i_1},\cdots,e_{i_k}) \sum_{\sigma\in\mathcal{S}_k} \varepsilon_\sigma v_{\sigma(1)}^{i_1}\cdots v_{\sigma(k)}^{i_k}.$$

Autrement dit, tout élément de $\Lambda_k(E)$ est de la forme

$$\varphi = \sum_{1\le i_1<\cdots<i_k\le n} a_{i_1\cdots i_k} e_{i_1}^*\wedge\cdots\wedge e_{i_k}^* \quad \text{avec} \quad a_{i_1\cdots i_k}\in\mathbb{R}.$$

On vérifie aisément que la décomposition ci-dessus est unique. on a donc prouvé la

Proposition 7.1.6 *Supposons que $1\le k\le n$. Alors $\Lambda_k(E)$ est un espace vectoriel de dimension $\binom{n}{k}=\frac{n!}{k!(n-k)!}$ et a pour base la famille $\left(e_{i_1}^*\wedge\cdots\wedge e_{i_k}^*\right)_{1\le i_1<\cdots<i_k\le n}$.*

Remarque : Dans le cas particulier $k=n$, on obtient $\dim\Lambda_k(E)=1$. On constate que l'application $e_1^*\wedge\cdots\wedge e_n^*$ n'est autre que l'application déterminant associée à la base $(e_1,\cdots,e_n)$.

7.1.3 Opérations sur $\Lambda_k(E)$

En plus de l'addition et de la multiplication par un scalaire, on peut munir l'ensemble des formes multi-linéaires alternées de nouvelles opérations : le produit extérieur et la transposée par une application linéaire.

L'opération de **produit extérieur** (que l'on notera $\wedge$) permet de construire un élément de $\Lambda_{k+\ell}(E)$ à partir d'un élément de $\Lambda_k(E)$ et d'un élément de $\Lambda_\ell(E)$. Donnons en la définition :

Définition 7.1.7 *Soit* $\omega \in \Lambda_k(E)$ *et* $\eta \in \Lambda_\ell(E)$. *On définit un élément* $\omega \wedge \eta$ *de* $\Lambda_{k+\ell}(E)$ *par*

$$\forall (v_1, \cdots, v_{k+\ell}) \in E^{k+\ell},\ \omega\wedge\eta(v_1, \cdots, v_{k+\ell}) = \sum_{\sigma \in \mathcal{S}_{k+\ell}} \varepsilon_\sigma \omega(v_{\sigma(1)}, \cdots, v_{\sigma(k)})\eta(v_{\sigma(k+1)}, \cdots, v_{\sigma(k+\ell)}).$$

Remarque : Nous laissons au lecteur le soin de vérifier que la notation $\wedge$ est cohérente avec celle que nous avons introduite pour définir les formes bilinéaires $e_i \wedge e_j$, puis (par récurrence) les formes k-linéaires $e_{i_1} \wedge \cdots \wedge e_{i_k}$.

À titre d'exercice, on pourra également montrer la proposition suivante :

Proposition 7.1.8 *Soit* $\omega \in \Lambda_k(E)$, $\eta \in \Lambda_\ell(E)$ *et* $\alpha \in \Lambda_m(E)$. *Alors on a :*

(i) $\forall (\lambda, \mu) \in \mathbb{R}^2,\ (\lambda\omega + \mu\eta) \wedge \alpha = \lambda(\omega \wedge \alpha) + \mu(\eta \wedge \alpha)$,

(ii) $(\omega \wedge \eta) \wedge \alpha = \omega \wedge (\eta \wedge \alpha)$,

(iii) $\omega \wedge \eta = (-1)^{k\ell}\, \eta \wedge \omega$.

Remarque : Si l'on se donne n formes linéaires $f_1, \cdots, f_n$ alors on a

$$\forall (x_1, \cdots, x_n) \in E^n, f_1 \wedge \cdots \wedge f_n(x_1, \cdots, x_n) = \det\big((f_i(x_j))_{1 \le i,j \le n}\big). \tag{7.1}$$

En effet, en reprenant la définition du produit extérieur, on trouve

$$f_1 \wedge \cdots \wedge f_n(x_1, \cdots, x_n) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon_\sigma f_1(x_{\sigma(1)}) \cdots f_n(x_{\sigma(n)}).$$

Dans le second membre, on reconnaît le déterminant de la matrice $(f_i(x_j))_{1 \le i,j \le n}$.

Passons maintenant à la définition de la *transposée par rapport à une application linéaire.* Avant d'aborder le cas général, rappelons la définition de transposée d'une application linéaire telle qu'elle est enseignée en deuxième année de licence :

Définition 7.1.9 *Soit* E *et* F *deux espaces vectoriels sur* $\mathbb{R}$, *et* $f \in L(E; F)$. *On appelle* **transposée** *de* f *l'application* ${}^T\!f$ *de* $L(F^*; E^*)$ *définie par* ${}^T\!f(u) = u \circ f$ *pour tout* $u \in F^*$.

Remarque : on a donc $\forall u \in F^*,\ \forall v \in E,\ [{}^T\!f(u)](v) = u(f(v))$.

Si l'on adopte le "langage" des applications multilinéaires alternées, l'opération de transposition définie ci-dessus permet de construire un élément de $\Lambda_1(E)$ à partir d'un élément de $\Lambda_1(F)$ et d'une application linéaire de E vers F.

Ce procédé peut se généraliser comme suit :

Définition 7.1.10 *Soit* $f \in L(E; F)$. *On appelle* k-**transposée** *de* f *l'application* ${}^T\!f$ *linéaire de* $\Lambda_k(F)$ *sur* $\Lambda_k(E)$ *définie par :*

$$\forall u \in \Lambda_k(F),\ \forall (v_1, \cdots, v_k) \in E^k,\ [{}^T\!f(u)](v_1, \cdots, v_k) \stackrel{\text{déf}}{=} u(f(v_1), \cdots, f(v_k)).$$

Attention : La notation ${}^T\!f$ ne tient pas compte de la dépendance en k. En pratique, cela ne pose pas de problème.

Proposition 7.1.11 *Pour tout* $\omega \in \Lambda_k(F)$, $\eta \in \Lambda_\ell(F)$ *et* $f \in L(E; F)$, *on a*

$${}^T\!f(\omega \wedge \eta) = {}^T\!f(\omega) \wedge {}^T\!f(\eta).$$

Preuve : Il suffit de calculer : on fixe $(v_1, \cdots, v_{k+\ell}) \in E^{k+\ell}$ et l'on a, par définition du produit extérieur et de la transposition :

$$\begin{aligned} {}^T\!f(\omega \wedge \eta)(v_1, \cdots, v_{k+\ell}) &= (\omega \wedge \eta)(f(v_1), \cdots, f(v_{k+\ell})), \\ &= \sum_{\sigma \in \mathcal{S}_{k+\ell}} \varepsilon_\sigma \, \omega\big(f(v_{\sigma(1)}), \cdots, f(v_{\sigma(k)})\big) \, \eta\big(f(v_{\sigma(k+1)}), \cdots, f(v_{\sigma(k+\ell)})\big). \end{aligned}$$

Par ailleurs, on a

$$\begin{aligned} {}^T\!f(\omega) \wedge {}^T\!f(\eta)(v_1, \cdots, v_{k+\ell}) &= \sum_{\sigma \in \mathcal{S}_{k+\ell}} \varepsilon_\sigma {}^T\!f(\omega)\big(v_{\sigma(1)}, \cdots, v_{\sigma(k)}\big) {}^T\!f(\eta)\big(v_{\sigma(k+1)}, \cdots, v_{\sigma(k+\ell)}\big), \\ &= \sum_{\sigma \in \mathcal{S}_{k+\ell}} \varepsilon_\sigma \, \omega\big(f(v_{\sigma(1)}), \cdots, f(v_{\sigma(k)})\big) \, \eta\big(f(v_{\sigma(k+1)}), \cdots, f(v_{\sigma(k+\ell)})\big), \end{aligned}$$

d'où le résultat. ■

7.2 Formes différentielles

Dans toute cette partie, U désigne un ouvert de $\mathbb{R}^n$ et V, un ouvert de $\mathbb{R}^p$. On munit $\mathbb{R}^n$ et $\mathbb{R}^p$ de leur base canonique.

7.2.1 Définitions

Définition 7.2.1 *Soit $k \in \mathbb{N}^*$. On appelle* **k-forme différentielle** *toute application définie sur un ouvert U de $\mathbb{R}^n$ et à valeurs dans $\Lambda_k(\mathbb{R}^n)$.*

Convention : Une 0-forme différentielle est une fonction de U dans $\mathbb{R}$.

Remarque : D'après la section précédente, toute forme k-linéaire alternée sur $\mathbb{R}^n$ peut s'écrire comme combinaison linéaire de k-formes élémentaires construites à l'aide de la base duale de la base canonique de $\mathbb{R}^n$. En reprenant les notations de la section 2.3 du chapitre 2, la base duale de $\mathbb{R}^n$ s'écrit $(dx_1, \cdots, dx_n)$. Donc toute k-forme différentielle définie sur l'ouvert U de $\mathbb{R}^n$ est de la forme :

$$\alpha = \sum_{1 \le i_1 < \cdots < i_k \le n} a_{i_1 \cdots i_k} dx_{i_1} \wedge \cdots \wedge dx_{i_k} \quad \text{avec} \quad a_{i_1 \cdots i_k} : U \to \mathbb{R}.$$

Sur cette formule, on voit que la k-forme α est continue (resp. différentiable) si et seulement si ses coefficients $a_{i_1 \cdots i_k}$ sont continus (resp. différentiables), etc.

Exemples :

1. Une 1-forme différentielle sur l'ouvert U de $\mathbb{R}^n$ s'écrit

$$\alpha = a_1 dx_1 + \cdots + a_n dx_n \quad \text{avec} \quad a_i : U \to \mathbb{R}.$$

Rappelons que si $f : U \to \mathbb{R}$ est différentiable, la différentielle df de f vérifie :

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i.$$

Donc df peut être identifiée à une 1-forme. Nous verrons plus loin que certaines 1-formes ne peuvent pas s'écrire comme différentielle d'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$.

Notons au passage que la différentielle de l'application $x \mapsto x_i$ n'est autre que la 1-forme constante égale à dx_i. La notation $(dx_1, \cdots, dx_n)$ pour désigner la base duale de $\mathbb{R}^n$ est donc cohérente !

2. Toute 2-forme différentielle définie sur l'ouvert U de $\mathbb{R}^n$ peut s'écrire
$$\omega = \sum_{1\le i<j\le n} a_{ij}\, dx_i \wedge dx_j \quad \text{avec} \quad a_{ij} : U \to \mathbb{R}.$$
Dans le cas $n = 2$, on note généralement (dx, dy) la base duale de $\mathbb{R}^2$. On voit alors que toutes les 2-formes sur $U \subset \mathbb{R}^2$ sont du type $\omega = a\, dx \wedge dy$ avec $a : U \to \mathbb{R}$.
3. Notons (dx, dy, dz) la base duale de $\mathbb{R}^3$. On constate que toutes les 3-formes différentielles définies sur un ouvert U de $\mathbb{R}^3$ sont du type
$$\eta = a\, dx \wedge dy \wedge dz \quad \text{avec} \quad a : U \to \mathbb{R}.$$

7.2.2 Changements de variables et transposition

a) Le cas d'une 1-forme différentielle

Considérons une application $f : U \subset \mathbb{R}^n \to \mathbb{R}$ différentiable. Nous avons vu dans la partie précédente que la différentielle de f pouvait être identifiée à la 1-forme
$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i.$$
Si l'on se donne un *changement de variable* c'est-à-dire une application $\phi : V \subset \mathbb{R}^p \to U \subset \mathbb{R}^n$ de classe C^1, nous avons établi dans la section 2.3 du chapitre 2 que l'application $F : V \to \mathbb{R}$ définie par $F \stackrel{\text{déf}}{=} f \circ \phi$ était également différentiable et que l'on pouvait écrire en notant $(dt_1, \cdots, dt_p)$ la base duale de $\mathbb{R}^p$:
$$\forall t \in V,\ dF(t) = \sum_{i=1}^{p} \frac{\partial F}{\partial t_i}(t) dt_i = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\phi(t))\, d\phi_j(t).$$
Nous souhaitons obtenir une expression explicite de l'application ϕ^* qui à la 1-forme df définie sur $U \subset \mathbb{R}^n$ associe la 1-forme dF définie sur $V \subset \mathbb{R}^p$. D'après le théorème de composition, on a
$$\forall t \in V,\ dF(t) = \big[\phi^*(df)\big](t) = df(\phi(t)) \circ d\phi(t) = {}^T\!\big(d\phi(t)\big)\big(df(\phi(t))\big).$$
Cette définition peut facilement se généraliser à toute 1-forme différentielle continue :

Définition 7.2.2 *Soit $\phi : V \to U$ une application de classe C^1 et α une 1-forme différentielle définie et continue sur U. On appelle* **transposée** *de α par l'application ϕ la 1-forme différentielle continue $\phi^*\alpha$ définie sur V par*
$$\forall t \in V,\ \phi^*\alpha(t) = \alpha(\phi(t)) \circ d\phi(t) = {}^T\!\big(d\phi(t)\big)\big(\alpha(\phi(t))\big).$$

Remarque : L'opération de transposition par l'application $\phi : V \to U$ permet donc d'associer à toute 1-forme définie sur l'ouvert U de $\mathbb{R}^n$, une 1-forme définie sur l'ouvert V de $\mathbb{R}^p$.

À l'aide de la définition de ϕ^*, on obtient facilement le résultat suivant :

Proposition 7.2.3 *La transposition par ϕ est une application linéaire de $\mathcal{C}\big(U; \Lambda_1(\mathbb{R}^n)\big)$ dans $\mathcal{C}\big(V; \Lambda_1(\mathbb{R}^p)\big)$. De plus, on a*
(i) $\forall i \in \{1, \cdots, n\},\ \phi^* dx_i = d\phi_i,$
(ii) $\forall \lambda \in \mathcal{C}(U; \mathbb{R}),\ \forall \alpha \in \mathcal{C}\big(U; \Lambda_1(\mathbb{R}^n)\big),\ \phi^*(\lambda\alpha) = \lambda\circ\phi\, \phi^*\alpha.$

Remarque : En combinant la linéarité de ϕ^* avec les deux propriétés ci-dessus, on retrouve que pour toute fonction $f \in \mathcal{C}^1(U; \mathbb{R}^n)$, on a
$$\phi^* df = dF \quad \text{avec} \quad F = f \circ \phi.$$

b) Le cas général

Pour définir la transposée d'une k-forme différentielle générale nous allons partir de la définition 7.2.2 qui garde un sens pour $k \geq 2$ si l'on interprète ${}^T d\phi(t)$ comme une k-transposée, c'est-à-dire comme l'application linéaire de $\Lambda_k(\mathbb{R}^n)$ dans $\Lambda_k(\mathbb{R}^p)$ introduite dans la définition 7.1.10.

Définition 7.2.4 *Supposons $\phi : V \to U$ de classe C^1. Soit α une k-forme différentielle continue sur U. On définit alors la k-forme différentielle $\phi^*\alpha$ continue sur V par*

$$\forall t \in V,\ \phi^*\alpha(t) = {}^T(d\phi(t))(\alpha(\phi(t)).$$

Remarque : Autrement dit, pour tout $t \in V$ et k-uplet $(v_1, \cdots, v_k)$ de vecteurs de $\mathbb{R}^p$, on a

$$\phi^*\alpha(t)(v_1, \cdots, v_k) = \alpha(\phi(t))(d\phi(t)(v_1), \cdots d\phi(t)(v_k)).$$

Proposition 7.2.5 *Pour tout $k \geq 1$, la transposition par ϕ est une application linéaire de $\mathcal{C}(U; \Lambda_k(\mathbb{R}^n))$ dans $\mathcal{C}(V; \Lambda_k(\mathbb{R}^p))$. De plus, pour toute fonction $\lambda : U \to \mathbb{R}$, k-forme différentielle α et ℓ-forme différentielle β définies sur U les identités suivantes sont vérifiées :*

(i) $\phi^*(\lambda\alpha) = \lambda\circ\phi\, \phi^*\alpha$,

(ii) $\phi^*(\alpha \wedge \beta) = \phi^*\alpha \wedge \phi^*\beta$.

Preuve : La première identité est une conséquence immédiate de la définition.

Pour prouver la deuxième identité, on écrit que, par définition de la transposée, on a pour tout $t \in V$,

$$\begin{cases} \phi^*(\alpha \wedge \beta)(t) &= {}^T d\phi(t)\Big[\alpha(\phi(t)) \wedge \beta(\phi(t))\Big], \\ \big(\phi^*\alpha \wedge \phi^*\beta\big)(t) &= \Big({}^T d\phi(t)\big(\alpha(\phi(t))\big)\Big) \wedge \Big({}^T d\phi(t)\big(\beta(\phi(t))\big)\Big). \end{cases}$$

La proposition 7.1.11 permet de conclure que les deux membres de droite sont égaux. ∎

c) Exemples

Calculer $\phi^*\alpha$ en fonction de ϕ et des coefficients de α est en général une tâche assez fastidieuse. Nous donnons ici quelques exemples de calcul de la transposée dans des cas simples.

- *Cas d'une 2-forme sur $\mathbb{R}^2$ et d'un changement de variable $\phi : \mathbb{R}^3 \to \mathbb{R}^2$.*

  En toute généralité, une 2-forme sur $\mathbb{R}^2$ peut s'écrire $\alpha = a\, dx \wedge dy$.

  Grâce à la proposition 7.2.5, on a

$$\begin{aligned} \phi^*\alpha &= a\circ\phi\, \phi^* dx \wedge \phi^* dy, \\ &= a\circ\phi \left(\tfrac{\partial\phi_1}{\partial s} ds + \tfrac{\partial\phi_1}{\partial t} dt + \tfrac{\partial\phi_1}{\partial u} du\right) \wedge \left(\tfrac{\partial\phi_2}{\partial s} ds + \tfrac{\partial\phi_2}{\partial t} dt + \tfrac{\partial\phi_2}{\partial u} du\right). \end{aligned}$$

  En développant puis en utilisant le fait que $ds \wedge ds = 0$, $ds \wedge du = -du \wedge ds$, etc., on obtient

$$\begin{aligned} \phi^*\alpha = a\circ\phi \bigg(\Big(\frac{\partial\phi_1}{\partial s}\frac{\partial\phi_2}{\partial t} - \frac{\partial\phi_2}{\partial s}\frac{\partial\phi_1}{\partial t}\Big) ds\wedge dt + \Big(\frac{\partial\phi_1}{\partial s}\frac{\partial\phi_2}{\partial u} - \frac{\partial\phi_2}{\partial s}\frac{\partial\phi_1}{\partial u}\Big) ds\wedge du \\ + \Big(\frac{\partial\phi_1}{\partial t}\frac{\partial\phi_2}{\partial u} - \frac{\partial\phi_2}{\partial t}\frac{\partial\phi_1}{\partial u}\Big) dt\wedge du\bigg). \end{aligned}$$

- *Cas d'une* 2*-forme sur* $\mathbb{R}^3$ *et d'un changement de variables* $\phi : \mathbb{R}^2 \to \mathbb{R}^3$.
  Toute 3-forme sur $\mathbb{R}^3$ est du type $\alpha = a dy \wedge dz + b dz \wedge dx + c dx \wedge dy$.
  D'après la proposition 7.2.5, on a

$$\begin{aligned}\phi^*\alpha &= a\circ\phi\, \phi^*dy \wedge \phi^*dz + b\circ\phi\, \phi^*dz \wedge \phi^*dx + c\circ\phi\, \phi^*dx \wedge \phi^*dy,\\ &= a\circ\phi \left(\tfrac{\partial\phi_2}{\partial s}ds + \tfrac{\partial\phi_2}{\partial t}dt\right)\wedge\left(\tfrac{\partial\phi_3}{\partial s}ds + \tfrac{\partial\phi_3}{\partial t}dt\right) + b\circ\phi \left(\tfrac{\partial\phi_3}{\partial s}ds + \tfrac{\partial\phi_3}{\partial t}dt\right)\wedge\left(\tfrac{\partial\phi_1}{\partial s}ds + \tfrac{\partial\phi_1}{\partial t}dt\right)\\ &\qquad + c\circ\phi \left(\tfrac{\partial\phi_1}{\partial s}ds + \tfrac{\partial\phi_1}{\partial t}dt\right)\wedge\left(\tfrac{\partial\phi_2}{\partial s}ds + \tfrac{\partial\phi_2}{\partial t}dt\right).\end{aligned}$$

  En développant puis en utilisant $ds \wedge ds = dt \wedge dt = 0$ et $dt \wedge ds = -ds \wedge dt$, on obtient

$$\phi^*\alpha = \left[a\circ\phi \left(\frac{\partial\phi_2}{\partial s}\frac{\partial\phi_3}{\partial t} - \frac{\partial\phi_3}{\partial s}\frac{\partial\phi_2}{\partial t}\right) + b\circ\phi \left(\frac{\partial\phi_3}{\partial s}\frac{\partial\phi_1}{\partial t} - \frac{\partial\phi_1}{\partial s}\frac{\partial\phi_3}{\partial t}\right) + c\circ\phi \left(\frac{\partial\phi_1}{\partial s}\frac{\partial\phi_2}{\partial t} - \frac{\partial\phi_2}{\partial s}\frac{\partial\phi_1}{\partial t}\right)\right] ds\wedge dt.$$

- *Cas d'une* 3*-forme sur* $\mathbb{R}^3$ *et d'un changement de variables* $\phi : \mathbb{R}^2 \to \mathbb{R}^3$.
  Dans ce cas $\phi^*\alpha$ est une 3-forme différentielle sur $\mathbb{R}^2$. Comme $3 > 2$, on peut affirmer sans calcul que $\phi^*\alpha = 0$.

Enfin, retenons le résultat ci-dessous qui est étroitement lié à la formule de changement de variable dans les intégrales multiples (voir le cours du second semestre) :

Proposition 7.2.6 *Si* $\phi : V \subset \mathbb{R}^n \to U \subset \mathbb{R}^n$ *est différentiable et* $\lambda : U \to \mathbb{R}$ *alors on a*

$$\phi^*\big(\lambda\, dx_1 \wedge \cdots \wedge dx_n\big) = \lambda\circ\phi\, \det(D\phi)\, \big(dt_1 \wedge \cdots \wedge dt_n\big).$$

Preuve : D'après la proposition 7.2.5, on a

$$\phi^*\big(\lambda\, dx_1 \wedge \cdots \wedge dx_n\big) = \lambda\circ\phi\, (\phi^*dx_1) \wedge \cdots \wedge (\phi^*dx_n).$$

Or d'après la proposition 7.2.3, on a $\phi^*dx_i = d\phi_i$ pour tout $i \in \{1, \cdots, n\}$. En utilisant (7.1) compte tenu de $d\phi_i(\epsilon_j) = \frac{\partial\phi_i}{\partial x_j}$ (où ϵ_j désigne le j-ième vecteur de la base canonique de $\mathbb{R}^n$), on obtient alors l'égalité souhaitée. ■

7.2.3 Formes exactes et formes fermées

Commençons par définir une nouvelle opération de différentiation : la différentielle extérieure.

Définition 7.2.7 *Soit* α *une* k*-forme différentielle différentiable sur l'ouvert* U *de* $\mathbb{R}^n$ *:*

$$\alpha = \sum_{1\le i_1<\cdots<i_k\le n} a_{i_1\cdots i_k} dx_{i_1} \wedge \cdots \wedge dx_{i_k}.$$

On définit alors la $(k+1)$*-forme différentielle* $d\alpha$ *par la formule :*

$$d\alpha \stackrel{\text{déf}}{=} \sum_{1\le i_1<\cdots<i_k\le n} da_{i_1\cdots i_k} \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k}.$$

La forme différentielle $d\alpha$ *est appelée* **différentielle extérieure** *de* α.

Remarque : Prendre la différentielle extérieure d'une 0-forme c'est-à-dire d'une fonction de plusieurs variables "ordinaire" revient à calculer la différentielle habituelle.

Attention : Dans le cas général, ne pas confondre la différentielle extérieure d'une k-forme (qui est une application de U dans $\Lambda_{k+1}(\mathbb{R}^n)$) avec la différentielle "ordinaire" (qui est une fonction de U dans $\mathcal{L}(\mathbb{R}^n; \Lambda_k(\mathbb{R}^n))$. Les deux notions se rejoignent uniquement si $k = 0$. Pour $k \ge 1$, différentielle extérieure et différentielle ne sont pas des objets de même nature et ne peuvent pas être mis en relation de façon simple !

Exemples

- *Différentielle extérieure d'une* 1 *forme sur* $\mathbb{R}^n$.
  Soit $\alpha = \sum_{j=1}^n a_j dx_j$. Comme $dx_j \wedge dx_i = -dx_j \wedge dx_i$ et $da_j = \displaystyle\sum_{i=1}^n \frac{\partial a_j}{\partial x_i} dx_i$, un calcul immédiat donne :

  (7.2) $$d\alpha = \sum_{1\le i<j\le n} \left(\frac{\partial a_j}{\partial x_i} - \frac{\partial a_i}{\partial x_j}\right) dx_i \wedge dx_j.$$

- *Différentielle extérieure d'une* 1 *forme sur* $\mathbb{R}^3$.
  Soit $\alpha = a\,dx + b\,dy + c\,dz$. En appliquant la formule (7.2), on obtient

  $$d\alpha = \Big(\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z}\Big) dy \wedge dz + \Big(\frac{\partial a}{\partial z} - \frac{\partial c}{\partial x}\Big) dz \wedge dx + \Big(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y}\Big) dx \wedge dy.$$

  Observons que si u désigne la fonction de $U \subset \mathbb{R}^3 \to \mathbb{R}^3$ de composantes (a, b, c), l'égalité ci-dessus peut se récrire à l'aide des composantes $\big((\mathrm{rot}\, u)_x, (\mathrm{rot}\, u)_y, (\mathrm{rot}\, u)_z\big)$ du **rotationnel** de u défini par

  $$\mathrm{rot}\, u = \nabla \wedge u = \begin{pmatrix} \frac{\partial c}{\partial y} - \frac{\partial b}{\partial z} \\ \\ \frac{\partial a}{\partial z} - \frac{\partial c}{\partial x} \\ \\ \frac{\partial b}{\partial x} - \frac{\partial a}{\partial y} \end{pmatrix}.$$

  En effet : on a $d\alpha = (\mathrm{rot}\, u)_x\, dy \wedge dz + (\mathrm{rot}\, u)_y\, dz \wedge dx + (\mathrm{rot}\, u)_z dx \wedge dy$.

- *Différentielle extérieure d'une* 2 *forme sur* $\mathbb{R}^3$.
  Soit $\alpha = a\, dy \wedge dz + b\, dz \wedge dx + c\, dx \wedge dy$. Par définition, on a

  $$d\alpha = da \wedge dy \wedge dz + db \wedge dz \wedge dx + dc \wedge dx \wedge dy.$$

  Après calcul, on obtient

  $$d\alpha = \Big(\frac{\partial a}{\partial x} + \frac{\partial b}{\partial y} + \frac{\partial c}{\partial z}\Big)\, dx \wedge dy \wedge dz.$$

  En posant encore $u = {}^T(a, b, c)$ puis en définissant la **divergence** $\mathrm{div}\, u$ de u par

  $$\mathrm{div}\, u = \nabla \cdot u = \frac{\partial a}{\partial x} + \frac{\partial b}{\partial y} + \frac{\partial c}{\partial z},$$

  on a alors $d\alpha = \mathrm{div}\, u\, dx \wedge dy \wedge dz$.

- *Différentielle extérieure d'une* 3 *forme sur* $\mathbb{R}^3$.
  La différentielle extérieure d'une 3-forme sur $\mathbb{R}^3$ est une 4-forme sur $\mathbb{R}^3$. On peut donc affirmer sans aucun calcul que la différentielle extérieure d'une 3-forme de $\mathbb{R}^3$ est nulle.

Définition 7.2.8 *On dit qu'une k-forme différentielle α est* **fermée** *si elle est différentiable et vérifie $d\alpha = 0$.*

Exemple : Les exemples ci-dessus permettent d'obtenir les résultats suivants :

- La 1-forme différentiable $\alpha = a\,dx + b\,dy + c\,dz$ est fermée si et seulement si les coefficients a, b et c vérifient :

  $$\frac{\partial c}{\partial y} = \frac{\partial b}{\partial z}, \quad \frac{\partial a}{\partial z} = \frac{\partial c}{\partial x} \quad \text{et} \quad \frac{\partial b}{\partial x} = \frac{\partial a}{\partial y}.$$

  En notant $u = {}^T(a, b, c)$, on peut donc affirmer que α est fermée si et seulement si $\mathrm{rot}\, u = 0$.

- La 2-forme différentiable $\alpha = a\,dy \wedge dz + b\,dz \wedge dx + c\,dx \wedge dy$ est fermée si et seulement si les coefficients a, b et c vérifient :

$$\frac{\partial a}{\partial x} + \frac{\partial b}{\partial y} + \frac{\partial c}{\partial z} = 0,$$

autrement dit si et seulement si $\operatorname{div} u = 0$.

Définition 7.2.9 *On dit qu'une $k+1$-forme α est* **exacte** *s'il existe une k forme différentielle ω telle que $\alpha = d\omega$.*

Exemples :

- La 1-forme différentiable $\alpha = a\,dx + b\,dy + c\,dz$ est exacte si et seulement si il existe une fonction différentiable $f : U \to \mathbb{R}$ telle que

$$a = \frac{\partial f}{\partial x}, \quad b = \frac{\partial f}{\partial y} \quad \text{et} \quad c = \frac{\partial f}{\partial z}.$$

Autrement dit, une 1-forme différentielle est exacte si et seulement si elle peut s'écrire comme la différentielle d'une fonction de U dans $\mathbb{R}$.

- La 2-forme $\alpha = a\,dy \wedge dz + b\,dz \wedge dx + c\,dx \wedge dy$ est exacte si et seulement si il existe trois fonctions f, g, h de U dans $\mathbb{R}$, différentiables telles que

$$a = \frac{\partial h}{\partial y} - \frac{\partial g}{\partial z}, \quad b = \frac{\partial f}{\partial z} - \frac{\partial h}{\partial x} \quad \text{et} \quad c = \frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}.$$

En posant encore $u = {}^T(a, b, c)$, cela revient à dire qu'il existe une fonction différentiable $v : U \to \mathbb{R}^3$ telle que $u = \operatorname{rot} v$.

Théorème 7.2.10 *Soit ω une k-forme différentielle à coefficients C^2. Alors on a $d(d\omega) = 0$.*

Preuve : Écrivons $\omega = \sum_{1 \le i_1 < \cdots < i_k \le n} \omega_{i_1 \cdots i_k}\, dx_{i_1} \wedge \cdots \wedge dx_{i_k}$ avec $\omega_{i_1 \cdots i_k}$ différentiable.

On a par définition de la différentielle extérieure,

$$d\omega = \sum_{1 \le i_1 < \cdots < i_k \le n} \sum_{i=1}^{n} \frac{\partial \omega_{i_1 \cdots i_k}}{\partial x_i}\, dx_i \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k}$$

puis

$$d(d\omega) = \sum_{1 \le i_1 < \cdots < i_k \le n} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 \omega_{i_1 \cdots i_k}}{\partial x_j \partial x_i}\, dx_i \wedge dx_j \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k}.$$

D'après le théorème de Schwarz, on a $\frac{\partial^2 \omega_{i_1 \cdots i_k}}{\partial x_j \partial x_i} = \frac{\partial^2 \omega_{i_1 \cdots i_k}}{\partial x_i \partial x_j}$. Comme $dx_i \wedge dx_j = -dx_j \wedge dx_i$, les termes de la double somme en i et j se compensent deux à deux, et l'on obtient bien $d(d\omega) = 0$. ■

Corollaire 7.2.11 *Toute forme différentielle exacte à coefficients C^1 est fermée.*

Preuve : Soit α une forme différentielle exacte. Il existe alors une forme différentielle ω à coefficients C^2 telle que $\alpha = d\omega$. On a donc d'après le théorème précédent $d\alpha = d(d\omega) = 0$. Donc α est fermée. ■

La réciproque du corollaire précédent est vraie sous des hypothèses topologiques portant sur le domaine de définition U de la forme différentielle. Retenons l'énoncé suivant :

Lemme 7.2.12 (de Poincaré) *Soit U un ouvert convexe*[1]*. Soit α une k-forme différentielle fermée et de classe C^1 sur U. Alors α est exacte.*

Preuve : Pour simplifier, nous nous limitons à la preuve dans le cas d'une 1-forme. Pour la démonstration dans le cas général, on pourra se reporter à [2] ou à [10]. Quitte à effectuer une translation, on peut toujours supposer que 0 est un point de U.

Notons $\alpha = \sum_{i=1}^n \alpha_i dx_i$ la 1-forme considérée. Rappelons que d'après la formule de Taylor avec reste intégral, on a pour toute fonction f de classe C^1 sur U,

$$\forall x \in U,\ f(x) = f(0) + \int_0^1 \sum_{i=1}^n x_i \frac{\partial f}{\partial x_i}(tx)\, dt.$$

Donc s'il existe une fonction f (que l'on peut prendre nulle en 0 sans perte de généralité) vérifiant $\alpha = df$, elle est nécessairement donnée par la formule :

$$\forall x \in U,\ f(x) = \int_0^1 \sum_{i=1}^n x_i \alpha_i(tx)\, dt.$$

Cette définition a bien un sens puisque tous les coefficients α_i sont continus et, par convexité de U, le segment $[0, x]$ est inclus dans U.

Il reste à vérifier que l'on a bien $df(x) = \alpha(x)$ pour tout $x \in U$. On calcule

$$df(x) = \sum_{j=1}^n \left[\int_0^1 \sum_{i=1}^n \left(tx_i \frac{\partial \alpha_i}{\partial x_j}(tx) + \delta_{ij}\alpha_i(tx) \right) dt \right] dx_j. \tag{7.3}$$

La condition $d\alpha = 0$ est équivalente à $\frac{\partial \alpha_i}{\partial x_j} = \frac{\partial \alpha_j}{\partial x_i}$ pour tout $(i,j) \in \{1, \cdots, n\}^2$. Par ailleurs, on constate que

$$\frac{\partial}{\partial t}\Big(\alpha_j(tx)\Big) = \sum_{i=1}^n x_i \frac{\partial \alpha_j}{\partial x_i}(tx).$$

En effectuant une intégration par parties, on obtient donc

$$\begin{aligned} \int_0^1 \sum_{i=1}^n tx_i \frac{\partial \alpha_i}{\partial x_j}(tx)\, dt &= \int_0^1 t \frac{\partial}{\partial t}\Big(\alpha_j(tx)\Big)\, dt, \\ &= \alpha_j(x) - \int_0^1 \alpha_j(tx)\, dt. \end{aligned}$$

En reportant ce résultat dans (7.3), on conclut que

$$df(x) = \sum_{j=1}^n \alpha_j(x)\, dx_j$$

comme souhaité. ■

Exemple : Soit α la 1-forme différentielle sur $\mathbb{R}^2$ définie par

$$\alpha(x,y) = (2x\cos y - y^2 \sin x)dx + (2y\cos x - x^2 \sin y)dy.$$

Un calcul aisé (exercice : le faire) donne $d\alpha = 0$. D'après le théorème ci-dessus, il doit donc exister une 0-forme (c'est-à-dire une fonction) telle que $\alpha = df$.

On vérifie en intégrant "à la main" que toutes les fonctions $f(x,y) = x^2\cos y + y^2 \cos x + C$ avec C réel arbitraire conviennent.

[1] Cette hypothèse peut être considérablement affaiblie : il suffit que U soit étoilé par rapport à un point ou même simplement connexe, voir [2] ou [10] pour plus de détails.

Corollaire 7.2.13 *Soit* U *un ouvert convexe de* $\mathbb{R}^3$ *et* $u : U \to \mathbb{R}^3$ *de classe* C^1*. Alors il existe une fonction* $f : U \to \mathbb{R}$ *de classe* C^2 *telle que* $u = \nabla f$ *si et seulement si* $\operatorname{rot} u = 0$.

Preuve : Un calcul immédiat montre que $u = \nabla f$ entraîne $\operatorname{rot} u = 0$.

Pour établir la réciproque, on note a, b et c les composantes de u et l'on pose $\alpha = a\,dx + b\,dy + c\,dz$. Sachant que $\operatorname{rot} u = 0$, un calcul précédent montre que $d\alpha = 0$. Le lemme de Poincaré permet de conclure à l'existence d'une fonction f (de classe C^2) telle que $\alpha = df$. Autrement dit, $u = \nabla f$. ■

Exercice : Soit U un ouvert convexe (ou même seulement simplement connexe) de $\mathbb{R}^3$ et $u : U \to \mathbb{R}^3$ de classe C^1. Montrer que $\operatorname{div} u = 0$ si et seulement si il existe $v : U \to \mathbb{R}^3$ de classe C^2 tel que $u = \operatorname{rot} v$.

Bibliographie

[1] P. Avez : Calcul différentiel, **Masson**.

[2] H. Cartan : Cours de calcul différentiel, **Hermann**.

[3] R. Danchin : Notes de cours d'analyse pour la préparation au CAPES, téléchargeable sur `http ://perso-math.univ-mlv.fr/users/danchin.raphael/`

[4] R. Danchin, R. Hadiji, S. Jaffard, S. Judée, E. Löcherbach, J. Printems et S. Seuret : Cours arithmétique et groupes, Licence première année, premier semestre, téléchargeable sur `http ://perso-math.univ-mlv.fr/users/printems.jacques/`

[5] J. Dieudonné : éléments d'analyse 1, **Gauthier-Villars**.

[6] J. Dixmier : Cours de mathématiques du premier cycle, deuxième année, **Gauthier-Villars**.

[7] J. Dixmier : Cours de topologie générale, **Presses Universitaires de France**.

[8] J.-M. Monier : Analyse 2, cours et 600 exercices corrigés, **Dunod**.

[9] F. Pham : Les différentielles, enseignement des mathématiques, **Masson**.

[10] M. Spivak : Calculus on manifolds, **Westview Press**.

[11] J.-C. Yoccoz : Calcul différentiel, **Publications de l'Université Paris Sud**.

Index